

Erwerb der lexikalischen Töne im Standardchinesischen durch deutsche Lerner

DISSERTATION

zur Erlangung des akademischen Grades

doctor philosophiae (Dr. phil)

im Fach Phonetik

angenommen von der Philosophischen Fakultät

Universität zu Köln

von Frau Tingting Brengelmann, M.A.

geboren am 30. 12. 1982 in Sichuan

Erstgutachterin: Prof. Dr. Martine Grice

Zweitgutachter: PD Dr. Stefan Baumann

Eingereicht am 06. April 2016

Datum der mündlichen Prüfung: 29. Juni 2016

Für Alexander

Danksagung

Die vorliegende Dissertation wäre nicht möglich ohne die großzügige Unterstützung und Inspiration vieler Menschen. Mein größter Dank gilt meiner Doktormutter Prof. Dr. Martine Grice, die immer an mich geglaubt und mich in der Forschung stets gefördert hat. Durch ihre Betreuung habe ich mir Wissen und Fähigkeiten angeeignet, die ich am Anfang der Promotion nicht für realistisch gehalten habe. Ich konnte im Lauf meiner Promotion jederzeit mit meinen Problemen, Sorgen und Fragen zu ihr kommen und fand immer ein offenes Ohr. Sie gab mir ständig Denkanstöße, neue Ideen und konstruktive Ratschläge, wenn ich vor einer schwierigen Aufgabe stand. Sie nahm sich immer Zeit, wenn ich ihre Hilfe gebraucht habe. Vielen Dank, Martine!

Ein großer Dank geht auch an Dr. Stefan Baumann. Von seinem Intonationsseminar bis zur Betreuung meiner Dissertation erlebte ich Stefan stets als einen präzisen und anspruchsvollen aber unglaublich freundlichen Menschen und Mentor, der mir Klarheit bei vielen fachlichen Konfusionen vermitteln konnte.

Ich möchte mich insbesondere bei Dr. Silvia Dahmen bedanken. Sie hat mein Wissen im Bereich Fremdspracheerwerb und Aussprachedidaktik erheblich vertieft. Ich habe sehr von ihrer reichhaltigen Erfahrung profitiert. Ihre ermutigenden Worte und tatkräftige Unterstützung waren für die vorliegende Dissertation von enormem Wert. Ich erinnere mich sehr gern an unseren lebhaften und interessanten Diskussionen.

Dr. Francesco Cangemi möchte ich ausdrücklich danken für seine brillanten Ideen und rigoroses analytisches Denken, die mir half, meine Daten sinnvoll und gründlich zu analysieren. Nicht zuletzt durch seine analytische Methodik habe ich sehr viel von ihm lernen können.

Ich möchte mich auch bei allen Kollegen im Institut für Linguistik – Phonetik der Universität zu Köln bedanken, die mich von Anfang fachlich und seelisch unterstützt haben. Ich habe von jedem von ihnen viel gelernt und bin sehr stolz ein Teil dieses tollen Teams zu sein. Ich danke Theo Klinker, der mir technisch bei meinen Aufnahmen immer zur Seite stand und mir überhaupt erst beigebracht hat, Aufnahmen selbst durchzuführen. Ich danke Christine Röhr, die mir geduldig gezeigt hat, wie man ein Online-Experiment aufbaut, zusätzlich immer wieder von meinem Jammern belästigt wurde und mich immer wieder motiviert hat. Ich danke Dr. Doris Mücke und Dr. Anne Hermes, die mir immer Halt gegeben haben, wenn ich manches Mal in Verzweiflung geraten bin.

Ein großer Dank gilt auch meinen Probanden, die an den Studien der Arbeit teilgenommen haben. Ohne ihre Hilfe gäbe es diese Arbeit nicht.

Zum Schluss und deswegen besonders herzlich möchte ich meinem Ehemann Sandro und meiner Familie danken, die immer hinter mir standen und mich bei allen unterstützt haben, was ich tat. Sie haben mich durch die sehr arbeitsintensive Zeit getragen und mir den Rücken freigehalten. Auch bei meinem kleinen Sohn Alexander möchte ich mich bedanken, der schon in seinem so jungen Alter Verständnis zeigte, wenn seine Mama am Schreibtisch saß, statt mit ihm auf den Spielplatz zu gehen. Ohne ihre uneingeschränkte Unterstützung wäre die Anfertigung dieser Dissertation nicht möglich gewesen.

Inhaltsverzeichnis

1 Einleitung.....	14
2 Linguistischer Hintergrund	18
2.1 Töne im Standardchinesischen	18
2.1.1 Töne des Standardchinesischen in Zitierform.....	19
2.1.2 Töne des Standardchinesischen im Kontext.....	27
2.2 Tonale Koartikulation	31
2.2.1 Tonale Koartikulation in ostasiatischen Tonsprachen	32
2.2.2 Tonale Koartikulation im SC.....	35
2.2.3 Ursprung und Einflussfaktoren der tonalen Koartikulation	37
2.2.4 Verschiedene Modelle zu tonaler Koartikulation.....	38
2.3 Weitere Aspekte der Prosodie im Standardchinesischen.....	41
2.3.1 Wortakzent.....	41
2.3.2 Funktionalität der Grundfrequenz.....	43
2.4 Deutsche Intonation	47
2.4.1 Wortakzent.....	49
2.4.2 Tonakzent.....	52
3 Theorien und Modelle im Zweitspracherwerb.....	57
3.1 Interlanguage – Lernaltersprache	57
3.2 Negativer Transfer.....	59
3.3 Phonologischer und Phonetischer Transfer.....	61
3.4 Einflussfaktoren beim Ausspracherwerb	63
3.4.1 Lernsituation: Zweitsprache vs. Fremdsprache	63
3.4.2 Alter des ersten Sprachkontakts.....	64
3.4.3 Dauer des Aufenthaltes im Zielsprachland	65
3.5 ChaF im Unterricht	65
3.6 Lehrbuchmangel.....	67
3.7 Studien über Chinesisch als Fremdsprache (ChaF)	69
4 Projektbeschreibung	76
4.1 Forschungsfragen und Hypothesen	76
4.2 Produktion und Perzeption der Töne	78
4.3 Methodik	81
4.4 Zugrundeliegende Theorien und Modelle des Spracherwerbs	82
4.5 Probanden	82
4.5.1 Muttersprachliche Kontrollgruppe.....	82
4.5.2 Deutsche Probanden.....	83
4.5.3 Die Evaluierung der fremdsprachlichen Kompetenz.....	85
5 Untersuchung der tonalen Koartikulation.....	87
5.1 Hypothesen	87
5.1.1 Tonale Alignierung im Standardchinesischen	88
5.1.2 Tonale Alignierung im Deutschen	91
5.2 Methodik	94
5.2.1 Testmaterial.....	95
5.2.2 Ablauf der Datenerhebung	96
5.2.3 F0-Extraktion und Normalisierung.....	97
5.2.4 Variabilitätsanalyse.....	98
5.3 Ergebnis.....	100
5.3.1 F0-Konturen	100
5.3.2 Statistische Analyse.....	105
5.3.3 Wirkungsbereich der koartikulatorischen Effekte	109

5.3.4	Assimilation vs. Dissimilation	112
5.3.5	Tonale Realisierung der Tonsequenz T3-T3	117
5.4	Tonale Realisierung der Zielwörter in Äußerungen.....	120
5.5	Zusammenfassung der Ergebnisse.....	122
6	Realisierung von Wörtern mit Neutralton	124
6.1	Forschungshypothese.....	124
6.1.1	Metrische Struktur.....	124
6.1.2	Tonale Realisierung.....	125
6.2	Methodik	131
6.2.1	Testmaterialien	131
6.2.2	Ablauf der Datenerhebung	132
6.3	Ergebnisse	134
6.3.1	Zielwörter realisiert in Zitierform	134
6.3.2	Zielwörter realisiert in einer Äußerung.....	142
6.4	Zusammenfassung der Ergebnisse.....	149
7	Zusammenfassung, Diskussion und Ausblick	150
7.1	Zusammenfassung der Ergebnisse.....	150
7.1.1	Ergebnisse im Bereich Tonerwerb	150
7.1.2	Ergebnisse im Bereich tonaler Koartikulation	151
7.2	Diskussionen der Ergebnisse.....	153
7.2.1	Tonale Koartikulation in der L2-Produktion.....	153
7.2.2	Tonale Realisierung des Neutraltons.....	156
7.2.3	Vorläufiges Modell des Tonerwerbs deutscher ChaF-Lerner	158
7.3	Didaktische Konsequenzen	160
7.3.1	Gewichtung des Aussprachetrainings	160
7.3.2	Expliziter Erwerb der phonetischen Realisierungen.....	160
7.3.3	Kontrastive Darstellung von L1 und L2.....	161
7.3.4	Konkrete Maßnahmen im ChaF-Unterricht.....	162
7.4	Forschungsd desiderata und Ausblick	163

Tabellen und Abbildungen

Tabellen

Tabelle 2.1:	Darstellung der Töne mit verschiedenen Beschreibungssystemen
Tabelle 2.2:	Ton 3 Variationen nach Chen (2000: 20).
Tabelle 2.3:	Verschiedene Prominenzstufen (Grice & Baumann 2007)
Tabelle 2.4:	Vergleich der prosodischen Systeme des SC und des Deutschen
Tabelle 3.1:	Zusammenfassung der Fehleranalyse nach Miracle (1989)
Tabelle 4.1:	Übersicht der Lerner-Informationen. DL1 – DL10 nehmen an der ersten und DL11 – DL16 an der zweiten Studie teil
Tabelle 4.2:	Neue HSK-Leitlinie der mündlichen Prüfung (adaptiert vom „ <i>Bulletin of HSK Test</i> “ 2011).
Tabelle 5.1:	Signifikanz der <i>Likelihood Ratio Tests</i> für die <i>mixed linear regression models</i> der Variabilität _{SEPARAT} des Offsets und Onsets.
Tabelle 5.2:	Ergebnisse der <i>mixed linear regression models</i> beim Vergleich der Werte der Messpunkte von Lernern und Muttersprachlern bei der Produktion der Tonsequenz T3-T3
Tabelle 6.1	Oberflächenkontur des Neutraltons nach T1, T2, T3 und T4, adaptiert von Shih (1987)

Tabelle 6.2: Ergebnisse der *mixed linear regression models* beim Vergleich der F0-Werte der Messpunkte zwischen Lernern und Muttersprachlern.

Tabelle 6.3: Ergebnisse der *mixed linear regression models* beim Vergleich der Werte der Messpunkte zwischen der Realisierung der Zielwörter in Äußerungen bei Lernern und Muttersprachlern.

Abbildungen

- Abbildung 2.1: Vier lexikalischen vollen Töne mit ihren F0-Konturen und die Darstellung in verschiedenen Transkriptionssystemen.
- Abbildung 2.2: Merkmalssystem von Yip (1980: 45)
- Abbildung 2.3: Typische Konturen zweisilbiger Wörter mit dem Neutralton auf der zweiten Silbe.
- Abbildung 2.4: Unterschiedliche Realisierungen des T3 in Zitierform von zwei Muttersprachlern.
- Abbildung 2.5: Phonetische Realisierung von T2 und Sandhi-T2
- Abbildung 2.6: Darstellung des *constraint*-basierten Modells der tonalen Koartikulation nach Flemming (2011).
- Abbildung 2.7: Globale Tonhöhenbewegungen in verschiedenen *tunes* nach Shen (1990a)
- Abbildung 2.8: F0-Konturen von Aussagesätzen (S) und Fragen (Q) mit Fokus in verschiedenen Positionen nach Xu (2005: 236).
- Abbildung 2.9: Beispiel für eine Realisierung des T4 mit einem größeren F0-Umfang in einer post-fokalen Position (Chen 2012: 112)
- Abbildung 2.10: Schematische Darstellung unterschiedlicher Typen der Hervorhebung im Deutschen, adaptiert von Terken & Hermes (2000)

- Abbildung 2.11: Sechs Tonakzente im Deutschen in GToBI
(www.gtobi.uni-koeln.de)
- Abbildung 2.12: GToBI-Annotation zum Tonakzent L*
(www.gtobi.uni-koeln.de)
- Abbildung 2.13: GToBI-Annotation zum Tonakzent L*+H
(www.gtobi.uni-koeln.de)
- Abbildung 2.14: Prosodische Funktionen und ihre intonatorischen Realisierungen. Grice & Baumann (2007)
- Abbildung 3.1: Schematische Darstellung der Bestandteile einer Lernersprache (Major 2001: 6)
- Abbildung 3.2: Das vorläufige Model des Tonerwerbs nach Yang (2011: 166)
- Abbildung 4.1: Vergleich zwischen Produktion (linke Spalte, a1 und a2) und Perzeption (rechte Spalte b1 und b2) bei Muttersprachlern des SC (Yang 2010: 98)
- Abbildung 5.1: Beispiel der progressiven Koartikulation anhand der Realisierung von T1 in P1 und P2 eines zweisilbigen Wortes (Xu 1997: 69 und 75).
- Abbildung 5.2.: Schematische Darstellung von hypothetischen Tontargets (gestrichelte Linien) und der oberflächlichen Kontur der Realisierung (durchgezogene Linien) nach Xu & Wang (2001: 321).

- Abbildung 5.3: Stilisierte Tonhöheverläufe, um „segmentale Verankerungen“ darzustellen, Ladd (2008: 175).
- Abbildung 5.4: Hypothetische Schemata der F0-Kontur von Tonsequenz T3-T1, realisiert von deutschen Lernern (DL) und Muttersprachlern (CM).
- Abbildung 5.5: Stilisierter Übergang von einem hohen zu einem tiefen Tontarget. (Flemming 2011)
- Abbildung 5.6: Darstellung der Berechnung der **Variabilität**_{SEPARAT} am Beispiel von T2 in P1 eines zweisilbigen Wortes.
- Abbildung 5.8: F0-Konturen von Zielsilben in P1 eines zweisilbigen Wortes. (a.) stellt die Produktion von Lernern und (b.) die Produktion von Muttersprachlern dar.
- Abbildung 5.9: F0-Konturen von Zielsilben in P2 eines zweisilbigen Wortes. (c.) stellt die Produktion von Lernern und (d.) die Produktion von Muttersprachlern dar.
- Abbildung 5.10: Koeffizienten der Variabilität im statistischen Modell am Offset (links) und am Onset (rechts) sowie der Interaktion von CONDITION.
- Abbildung 5.11: Die Koeffizienten der Variabilität in statistischen Modellen für einzelne Töne am Offset (links) und am Onset (rechts) in jedem Teilbild sowie der Interaktion von CONDITION.
- Abbildung 5.12: Wirkungskraft der Variable CONDITION auf die Variabilität der Messpunkte von Tönen in P1 (linke

Spalte) und P2 (rechte Spalte) bei Lernern (rot) und Muttersprachlern (blau).

Abbildung 5.13: Regressiver Effekt auf den Mittelwert der letzten zwei Messpunkte einer Silbe in P1.

Abbildung 5.14: Progressiver Effekt auf den Mittelwert der ersten zwei Messpunkte von Silben in P2.

Abbildung 5.15: Tonale Realisierung der T3-T3 Sequenz mit unterschiedlichem segmentalem Inhalt und drei Wiederholungen von Lernern (rot) und Muttersprachlern (blau).

Abbildung 5.16: Obere und untere Vertrauensgrenze der Realisierung von Tonsequenz T3-T3 in Zitierform von Lernern (links) und Muttersprachlern (rechts).

Abbildung 5.17: Beispiel für Konturfehler in der Produktion von Zielwörtern in Zitierform und in Äußerungen bei DL4 und DL5.

Abbildung 5.18: Beispiele für stark verlängerte Pausen zwischen zwei Silben in den Nonsens-Zielwörtern „wámĩ“ (linke Spalte) und „wúmò“ (rechte Spalte).

Abbildung 6.1: F0-Konturen von Äußerungen, in denen die erste Silbe einen der vier vollen Töne und die letzte Silbe den T4 trägt (Chen & Xu 2006).

- Abbildung 6.2: Die F0-Kontur der Tonsequenzen T1-T0 (*wān le* – „gebogen“) und T4-T0 (*màn le* – „langsam geworden“) sowie die Realisierung von H*+L- im Deutschen.
- Abbildung 6.3: Vier Zielwörter.
- Abbildung 6.4: Beispiel einer Bildsequenz mit dem Zielwort „*hóuzi*“ (,der Affe‘).
- Abbildung 6.5: Individuelle Darstellung der Realisierung von Wörtern mit Neutralton in Zitierform durch Lerner (linke Spalte) und Muttersprachler (rechte Spalte).
- Abbildung 6.6: Die obere und untere Vertrauensgrenze der Realisierung verschiedener Tonsequenzen in Zitierform bei verschiedenen Sprechergruppen.
- Abbildung 6.7: Dauerverhältnis zwischen der Silbe mit vollem Ton und Neutralton in Zitierform.
- Abbildung 6.8: Konturen zweisilbiger Wörter mit Neutralton von Lernern (rot) und Muttersprachlern (blau) in Äußerungen realisiert.
- Abbildung 6.9: Vertrauensgrenzen für Realisierungen der in Äußerungen produzierten Zielwörter Rot steht für Lerner und Blau für Muttersprachler.
- Abbildung 6.10: Das Dauerverhältnis zwischen den zwei Silben in Wörtern mit Neutralton.

- Abbildung 7.1: Beispiel für den Tonakzent L*+H
(www.gtobi.uni-koeln.de)
- Abbildung 7.2: Vorläufiges Modell zum Tonerwerb deutscher ChaF-Lerner
- Abbildung 7.3: F0-Konturen der Wörter 不 (*bù*, 'nicht') und 喝 (*hē*, 'trinken') produziert von einem Muttersprachler (in a.) sowie einem Lerner vor (b.) und nach (c.) einem einwöchigen audio-visuellen Training.

1 Einleitung

Die vorliegende Arbeit befasst sich mit den Problemen deutscher Lerner¹ beim Erwerb der lexikalischen Töne im Standardchinesischen. Es wird die Hypothese aufgestellt, dass die grundlegenden Mechanismen zur phonetischen Realisierung von Tonsequenzen durch deutsche Lerner aufgrund eines ungenügenden expliziten Trainings auf verschiedenen tontragenden Domänen entscheidend von denen chinesischer Muttersprachler abweichen.

Für einen Fremdsprachen-Lerner kann eine schlechte Aussprache zu beträchtlichen Kommunikationsschwierigkeiten führen (Mennen 2007). Viele Studien zeigen, dass Fehler auf der prosodischen Ebene einen destruktiveren Einfluss auf die Verständlichkeit der L2²-Produktion haben als Fehler auf der segmentalen Ebene (Anderson-Hsieh et.al. 1992; Johansson 1978; Nash 1972; Munro & Derwing 1995; Trofimovitch & Baker 2006). Innerhalb des breiten Feldes des Prosodie-Erwerbs stellt die Aneignung lexikalischer Töne eine außerordentlich große Herausforderung für erwachsene Lerner dar, deren Muttersprache³ keine Tonsprache ist (Orie 2006; Kiriloff 1969; Bluhme & Burr 1971; Shen 1989, unter anderen). Diese Schwierigkeiten können Lerner von Anfang an begleiten und führen in vielen Fällen nach ein oder zwei Jahren des Lernens zur Aufgabe (Orten 2013).

Ein Fünftel der Menschheit spricht Chinesisch als Muttersprache, doch nur einige tausend Menschen nichtchinesischer Abstammung beherrschen diese Sprache auf höherem Niveau (Bausch 2003: S. 513). Zum einen liegt dies daran, dass sich China bis in die 80er Jahre des 20. Jahrhunderts vom Rest der Welt abgeschottet hat, wodurch sich Standardchinesisch als Fremdsprache, im Gegensatz zu Sprachen wie Spanisch oder Englisch, erst sehr spät zu einem eigenständigen Fach entwickelte. Zum anderen liegt die schlechte Qualität des Tonerwerbs auch an

¹ Die Form „Lerner“ wird in der vorliegenden Arbeit stellvertretend für männliche sowie weibliche Lernende verwendet.

² „L2“ steht in der vorliegenden Arbeit für eine Fremdsprache/Zweitsprache.

³ Unter „Muttersprache“ versteht man die erste im Leben erworbene Sprache.

unzureichender Grundlagenforschung und ungeeignetem Unterricht.

Die Aussprachekompetenz, vor allem das prosodische Training, wird im ChaF⁴-Unterricht weitgehend vernachlässigt. Töne werden in der Regel ausschließlich in ihrer Zitierform vorgestellt und geübt. Eine explizite phonetische Übung zur tonalen Realisierung auf größeren tontragenden Domänen erfolgt üblicherweise nicht. Viele Faktoren spielen dabei eine Rolle, wie zum Beispiel die begrenzte oder sogar fehlerhafte Aufklärung über die chinesische Prosodie in den Lehrmaterialien und der kurze im Curriculum vorgesehene Zeitraum für das Aussprachetraining in den meisten ChaF-Kursen (Orton 2013).

Aufgrund der zunehmenden wirtschaftlichen und politischen Bedeutung Chinas beginnen immer mehr Deutsche, Standardchinesisch als Fremdsprache zu erwerben. Es besteht daher dringender Bedarf an einer verbesserten Grundlagenforschung und daraus entwickeltem Lehrmaterial. In den letzten Jahren hat es zahlreiche Studien zum Erwerb der lexikalischen Töne im Standardchinesischen gegeben (Chen 1997, 2000; Miracle 1989; Shen 1989; Sun 1998; White 1981; Zhang 2007, unter anderen). Diese Arbeiten beschäftigen sich hauptsächlich mit englischsprachigen Lernern. Die vorliegende Arbeit ist die erste systematische Studie zur Tonproduktion deutscher Lerner im Standardchinesischen.

Zudem behandelt die vorliegende Arbeit zwei Themen in der Tonerwerbsforschung, die noch nicht untersucht sind: Die tonale Koartikulation und die Realisierung des Neutraltons in der L2-Tonproduktion. Die Untersuchungen dazu werden mit messphonetischen Methoden durchgeführt und ausgewertet. Dabei werden tonale Realisierungen zweisilbiger Zielwörter von deutschen ChaF-Lernern und chinesischen Muttersprachlern aufgenommen und anhand gezielt aufgestellter Hypothesen verglichen und überprüft. Die Ergebnisse der Untersuchungen sollen Aufschluss darüber geben, welche Mechanismen deutsche

⁴ „ChaF“ steht für „Chinesisch als Fremdsprache“.

Lerner bei der Tonproduktion verwenden und welche Rolle der ChaF-Unterricht dabei spielt.

Die „kommunikative Wende“ in der Fremdsprachendidaktik hat die Interessensschwerpunkte vom möglichen Erreichen einer muttersprachlichen Aussprache hin zu kommunikativer Effizienz und Verständigungsfähigkeit in der Zielsprache verschoben (Missaglia 2001). Durch die ausschließlich pragmatische und handlungsorientierte Ausrichtung dieses kommunikativen Ansatzes wurde das Aussprachetraining der Prosodie der Zielsprache vernachlässigt (Missaglia 2001).

Die Ergebnisse der vorliegenden Arbeit plädieren für die Notwendigkeit der expliziten phonetischen Übung im Bereich Ausspracheerwerb. Insbesondere für den Erwerb der Töne im Standardchinesischen sollte sich die explizite phonetische Übung nicht nur auf Töne in Zitierform beschränken, sondern auf alle tontragenden Domänen ausgeweitet werden, wie zum Beispiel auf Wörter oder Äußerungen.

Diese Arbeit ist wie folgt aufgebaut:

In Kapitel 2 werden die linguistischen Theorien und Modelle des Standardchinesischen und Deutschen vorgestellt. Zuerst wird das tonale System im Standardchinesischen ausführlich beschrieben. In diesem Zusammenhang werden die linguistischen Eigenschaften des Neutraltons und der tonalen Koartikulation dargestellt, die für die empirischen Studien der vorliegenden Arbeit relevant sind. Dann wird die prosodische Struktur im Deutschen erläutert und zum Schluss der des Standardchinesischen gegenübergestellt.

In Kapitel 3 werden wichtige Theorien im Bereich Fremdspracherwerb für die vorliegende Arbeit vorgestellt, wie die *Interlanguage Hypothesis* oder die Theorie des Transfers. Einflussfaktoren im Ausspracheerwerb werden beschrieben, z.B. die Lernsituation und lernerspezifische Faktoren wie das Alter des ersten Sprachkontakts. Anschließend werden die Unterrichtspraxis und die aktuellen

Lehrmaterialien im Bereich ChaF beschrieben. Einige ausgewählte Studien zum Tonerwerb im Standardchinesischen werden danach vorgestellt. Dabei wird die Notwendigkeit der Untersuchung der grundlegenden Mechanismen der Lerner bei der Tonproduktion dargelegt.

Kapitel 4 ist die Überleitung vom theoretischen zum empirischen Teil der Arbeit. Zuerst werden die Hypothesen der vorliegenden Arbeit vorgestellt. Anschließend werden Aufbau und Methodik der empirischen Untersuchungen sowie die Theorien und Modelle des Spracherwerbs beschrieben, auf denen die vorliegende Arbeit beruht. Zum Schluss folgt eine kurze Vorstellung der muttersprachlichen und deutschen Probanden.

Mit Kapitel 5 beginnen die empirischen Untersuchungen der Arbeit. Hier werden tonale Realisierungen zweisilbiger Wörter mit zwei vollen Tönen und in Kapitel 6 zweisilbiger Wörter mit einem Neutralton untersucht. In diesen zwei Kapiteln werden nach erfolgter Aufstellung der Hypothesen die Testmaterialien, der Ablauf der Datenerhebung und die analytischen Ansätze beschrieben. Im Anschluss werden die Ergebnisse der Datenanalyse präsentiert.

Kapitel 7 fasst die Ergebnisse der beiden empirischen Studien zusammen. Diese werden hinsichtlich linguistischer und didaktischer Aspekte diskutiert. Die Hypothesen aus Kapitel IV werden überprüft. Anschließend werden didaktische Konsequenzen aus den Ergebnissen der vorliegenden Arbeiten beschrieben und zum Schluss noch offene Fragen und mögliche zukünftige Forschungsvorhaben vorgestellt.

2 Linguistischer Hintergrund

In diesem Kapitel werden die linguistischen Theorien und Modelle des Standardchinesischen und Deutschen vorgestellt. Zuerst wird das tonale System im Standardchinesischen ausführlich beschrieben und anschließend der prosodischen Struktur der deutschen Sprache gegenübergestellt.

2.1 Töne im Standardchinesischen

Die Chinesischen Sprachen können in die neun Dialektfamilien Wu, Min, Yue, Xiang, Gan, Jing Yu, Hui Yu, Hakka und Mandarin unterteilt werden. Sie werden auf dem chinesischen Festland, in Taiwan, Hong Kong, Malaysia, Singapur, Vietnam und in den Vereinigten Staaten von Amerika gesprochen. Die Familie Mandarin ist die größte mit ungefähr 70% der Sprecher. Standardchinesisch - auch *Putonghua* genannt - ist die Amtssprache in China (Duanmu 2000; He 2010; Yang 2011), dessen Definition und Normen in der Nationalen Reformkonferenz für Sprache in Oktober 1955 festgelegt wurden:

„普通话就是以北京语音为标准音, 以北方话为基础方言, 以典范的现代白话文著作为语法规范的现代汉民族共同语。“

[*Putonghua* ist die gemeinsame Sprache des modernen Han Volks, deren Aussprache der Aussprache des Dialekts von Beijing entspricht, deren Sprachverwendung des Mandarins sie als Standard übernimmt und deren Grammatik aus den modernen chinesischen klassischen Werken entnommen wird.]

(Ma, X.B. & Y. Zhao 2002: 1, meine Übersetzung)

1958 wird *Hanyu Pinyin*, eine Romanisierung des Standardchinesischen, von der ersten Volksversammlung in China erlassen. Es ist ein phonetisches Transkriptionssystem, das auf dem lateinischen Alphabet basiert und für westeuropäische Lerner die Basis der Aussprache des Standardchinesischen bildet.

Das tonale System des Standardchinesischen ist per Definition mit der Aussprache des Dialekts von Beijing identisch. Deshalb wurden nur Muttersprachler, die in Beijing geboren und aufgewachsen sind, als Sprecher der Kontrollgruppen der in der vorliegenden Arbeit durchgeführten Experimente rekrutiert.

Es gibt jedoch Unterschiede zwischen diesen zwei sprachlichen Varianten. Beispiel (2.1) a zeigt die Aussprache im Standardchinesischen, während (2.1) b die entsprechende Realisierung im Dialekt von Beijing darstellt.

(2.1)

- | | | | |
|----|---------------------------|-----------------|-------------------------|
| a. | <i>dǎi</i> | <i>xiǎo tōu</i> | |
| | fangen | klein stehlen | |
| | „den Dieb <i>fangen</i> “ | | im Standardchinesischen |
| | | | |
| b. | <i>děi</i> | <i>xiǎo tōu</i> | |
| | fangen | klein stehlen | |
| | „den Dieb <i>fangen</i> “ | | im Dialekt von Beijing |

(Lin, T. 2001: 23)

Fälle dieser Art treten meistens in informellen Sprechsituationen auf. Um derartige Fälle dennoch zu vermeiden, werden in der vorliegenden Arbeit nur Silben ausgewählt, die keine Abweichungen im Dialekt von Beijing aufweisen.

2.1.1 Töne des Standardchinesischen in Zitierform

Das Standardchinesisch (im Folgenden: SC) hat vier distinktive lexikalische Töne sowie einen Neutralton. Unter einem Ton versteht man einen Tonhöhenverlauf⁵, der lexikalisch distinktiv ist (Wiese 1988). Das heißt, zwei Silben mit identischen segmentalen Komponenten können allein durch die Töne ein Minimalpaar bilden.

⁵ „Tonhöhenverlauf“ wird in der vorliegenden Arbeit auch als „Kontur“ bezeichnet.

Die lexikalischen Töne und der Neutralton werden innerhalb des *Hanyu Pinyin* Systems wie in Beispiel (2.2) durch diakritische Zeichen bzw. deren Abwesenheit auf einem Vokal innerhalb der Silbe dargestellt. In der vorliegenden Arbeit werden Töne durch die Ziffern 0 bis 4 dargestellt.

(2.2)	Ton 1 (T1)	mā	„Mutter“
	Ton 2 (T2)	má	„Hanf“
	Ton 3 (T3)	mǎ	„Pferd“
	Ton 4 (T4)	mà	„schimpfen“
	Ton 0 (T0)	ma	Fragenpartikel

2.1.1.1 Die vier vollen⁶ lexikalischen Töne

Abbildung 2.1 stellt die vier vollen lexikalischen Töne des SC dar: gezeigt wird (von oben nach unten) der akustische Tonhöhenverlauf, die autosegmentale phonologische Repräsentation, bestehend aus abstrakten Hoch- und Tieftönen (H und L) und schließlich die Transliteration anhand der Silbe „ma“ in *Hanyu Pinyin*.

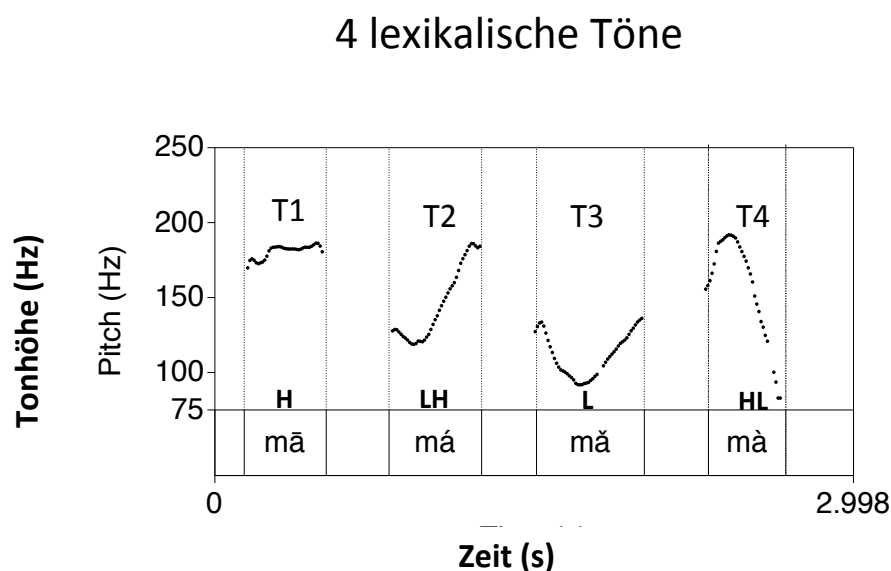


Abbildung 2.1: Vier lexikalischen vollen Töne mit ihren F0-Konturen und die Darstellung in verschiedenen Transkriptionssystemen.

⁶ Zur Abgrenzung vom Neutralton werden T1, T2, T3 und T4 in der vorliegenden Arbeit als „volle Töne“ bezeichnet.

T1 beginnt im oberen Bereich des Stimmumfangs und bleibt bis zum Ende der Silbe oben. T2 fängt im mittleren Bereich an und steigt stetig bis zum Ende der Silbe. T3 verläuft zum größten Teil im unteren Bereich des Stimmumfangs. Er startet tief, sinkt noch weiter bis zur untersten Grenze und steigt wieder nach oben in den mittleren Bereich. Die Variation von T3 im SC wird viel in der Forschung diskutiert und in Abschnitt 2.1.2.1 explizit erläutert. Der Onset-Wert von T4 ist der höchste unter allen Tönen im Stimmumfang eines Sprechers. Dieser Ton fällt steil von dem höchsten Punkt nach unten bis zum Silbenende. Diese typischen Tonhöheverläufe der Töne im SC wurden auch in akustischen Studien bestätigt (z.B. Ho 1976; Howie 1976, Tseng 1981).

In der Forschung existieren zwei Ansätze, die das tonale System des SC beschreiben. Chao Yuen-Ren (1930, 1968) war der erste Sprachwissenschaftler, der versucht hat, die chinesische Aussprache systematisch zu beschreiben. Er entwickelte 1930 ein Beschreibungssystem, die „Fünferskala“ (nach Wiese 1988), in der „1“ den niedrigsten Wert des Stimmumfangs eines Sprechers und „5“ den höchsten Wert repräsentiert. Tabelle 2.1 fasst verschiedene Beschreibungssysteme der Töne zusammen:

Töne	Konturen	Chao-Skala	Beispiele	Dt. Übersetzung
T1	flach	55	bā 八	„acht“
T2	steigend	35	bá 拔	„ausreißen“
T3	fallend- steigend	214	bǎ 把	„Knauf“
T4	fallend	51	bà 爸	„Vater“
T0 ⁷	mittig	3	ba 吧	Partikel

Tabelle 2.1: Darstellung der Töne mit verschiedenen Beschreibungssystemen. Adaptiert von Peng et al. (2005: 235)

⁷ In Abschnitt 2.1.1.2. wird der Neutralton detailliert beschrieben.

Der Status der Fünferskala von Chao ist intrinsisch vage, da die Unterteilung in fünf Ebenen nicht auf phonologischen Prinzipien beruht. Zum einen stellt diese Skala eine phonologische Abstraktion der relativen Tonhöhenverläufe dar und zum anderen versucht sie, phonetische Details der Töne zu erfassen (Duanmu 2000). Diese numerische Skala besitzt jedoch den Vorteil, dass man die inter-dialektalen und individuellen Variationen der tonalen Realisierung adäquat vergleichen kann, wie z.B. bei einem Vergleich der Tonproduktion zwischen Lernern und Muttersprachlern.

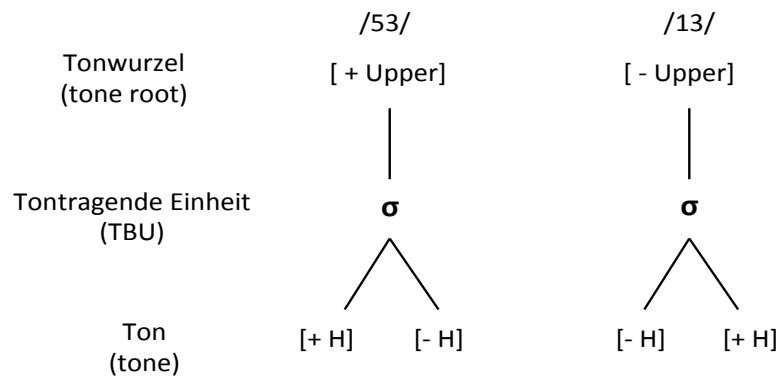
Anders als Chao geht Yip (1980) vom Modell der Autosegmentalen Phonologie aus und argumentiert für eine neue, merkmalsbasierte Analyse der chinesischen Töne. Yip schlägt ein Merkmalssystem (*feature system*) vor, in dem sie mithilfe zweier binärer Merkmale vier hierarchische Ebenen definiert, wie in Abbildung 2.2 demonstriert.

<u>Register</u>	<u>Tone</u>
+ Upper	+ High (H)
	- High (L)
- Upper	+ High (H)
	- High (L)

Abbildung 2.2: Merkmalssystem von Yip (1980: 45)

Das erste und gleichzeitig dominierende Merkmal in diesem System ist das **Registermerkmal**, das den Stimmumfang in zwei Hälften unterteilt. Das zweite Merkmal ist das **Tonmerkmal**, das einen H- und einen L-Ton beinhaltet. Yip setzt voraus, dass das Registermerkmal innerhalb eines Tons unverändert bleiben muss. Das Tonmerkmal kann wiederum innerhalb eines Tons alternieren, was die Konstruktion eines Konturtons ermöglicht. Demnach lässt sich ein Konturton als Verzweigungsknoten mit zwei terminalen Knoten mit gegensätzlichen tonalen Werten wie in Beispiel (2.3) darstellen:

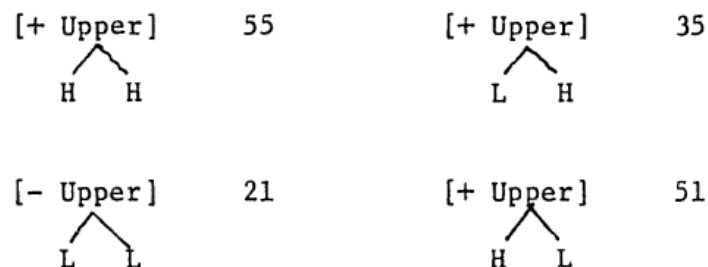
(2.3)



(Adaptiert von Yip 2002: 49)

Nach Yip (1980) lassen sich die vier vollen Töne im SC wie in Beispiel (2.4) darstellen:

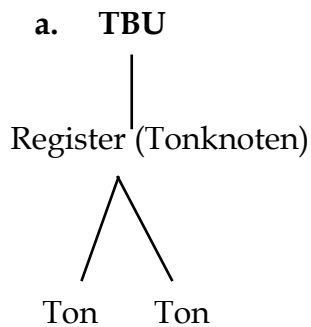
(2.4)



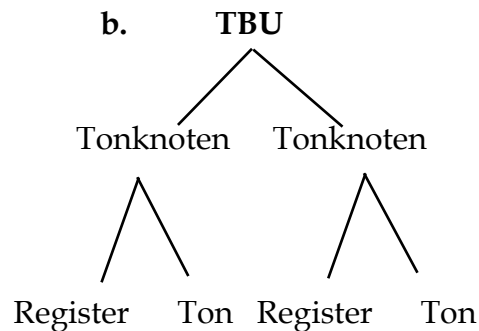
Yip (1980: 283)

Ähnliche merkmalsbasierte geometrische Bäume entwickelten Duanmu (1990) und Bao (1990). Alle drei Autoren unterteilen ihre Beschreibung in Register- und Tonmerkmale mit dem Unterschied in der Hierarchie der beiden Merkmale, wie in Beispiel (2.5) darstellt (Yip 2002: 53, Yang 2010: 33):

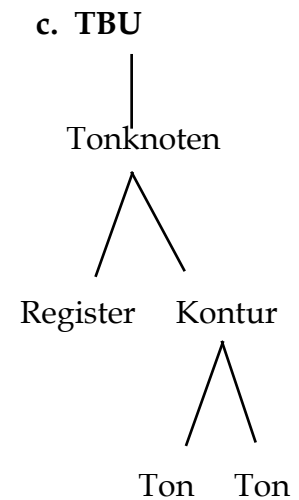
(2.5)



(Yip 1989)



(Duanmu 1990)



(Bao 1990)

In den Modellen (b) und (c) kann das Merkmal „Register“ auf mehrere Positionen verteilt werden, weil dieses dem Tonknoten untergeordnet ist. In Modell (a) dagegen ist das Merkmal Register selbst der Tonknoten, sodass dieses nicht mehr unterteilt werden kann.

Mit obigen Modellen können Töne innerhalb einer ähnlichen Struktur analysiert werden, auch wenn dabei die phonetische Variation in der tonalen Produktion nicht erfasst wird.

2.1.1.2 Der Neutralton

Der Neutralton im SC wird konventionell als ein Ton ohne phonologische Spezifikation beschrieben. Dieser hat eine eingeschränkte Verteilung, d.h., Silben mit Neutralton muss eine Silbe mit einem vollen Ton vorausgehen und sie dürfen nicht am Anfang eines prosodischen Wortes stehen (Yip 2002). Die vollen Töne kommen nur bei akzentuierten Silben vor und können bei nicht-akzentuierten Silben „neutralisiert“ werden. Shen (1992) unterteilt alle Silben mit Neutralton daher in intrinsisch tonlose Silben und von den vollen Tönen hergeleitete tonlose

Silben. Bei einem intrinsischen Neutralton handelt es sich meist um Funktionsmorpheme im SC (Li 1981), wie in Beispiel (2.6) auflistet:

- (2.6)
- | | |
|-----------|-------------------------------|
| <i>de</i> | Genitiv-Marker |
| <i>le</i> | Aspekt-Partikel/Modalpartikel |
| <i>ma</i> | Fragepartikel |
| <i>zi</i> | Nominalsuffix |

Der von vollen Tönen hergeleitete Neutralton wird „neutralisiert“. Aus diesem Grund haben Silben mit dieser Art von Neutralton ebenfalls eine Variante mit einem vollen Ton, wie in Beispiel (2.7):

- (2.7)
- | | |
|---|---|
| <p>a.</p> <p><i>dà yì</i></p> <p>groß Bedeutung</p> <p>„Hauptinhalt“</p>
<p><i>dà yì</i></p> <p>groß Bedeutung</p> <p>„achtlos“</p> | <p>b.</p> <p><i>dōng xī</i></p> <p>Ost West</p> <p>„Ost und West“</p>
<p><i>dōng xī</i></p> <p>Ost West</p> <p>„Ding“</p> |
|---|---|

Der Status des Neutraltons ist in der Forschung jedoch umstritten. Abbildung 2.3 zeigt die typischen Konturen zweisilbiger Wörter, bei denen der Neutralton jeweils einem der vier vollen Töne folgt.

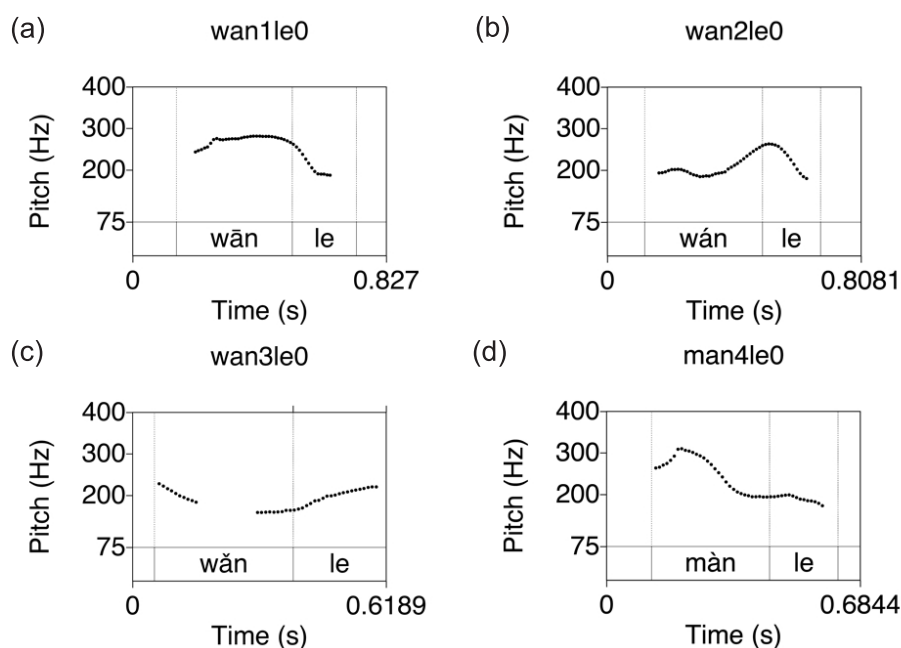


Abbildung 2.3: Typische Konturen zweisilbiger Wörter mit dem Neutralton auf der zweiten Silbe jeweils nach T1 (*wān le* – „gebogen“), T2 (*wán le* – „zum Ende gekommen“), T3 (*wǎn le* – „spät geworden“) und T4 (*màn le* – „langsam geworden“). Die Audioquelle wurden aus Peng et al. (2005) entnommen.

Es gibt verschiedene wissenschaftliche Ansätze, die Oberflächenkonturen von Silben mit Neutralton zu erklären. Yip (1980; 2002) und Shen (1992) sind der Meinung, dass die Oberflächenkontur des Neutraltons von dem vorausgehenden Ton durch Tonausbreitung (*tonal spreading*) bestimmt wird. Allerdings müsste die Kontur des Neutraltons nach T2 einen hohen Offset-Wert haben, was laut Abbildung 2.3 (b) nicht der Fall ist. Dagegen schlagen Shih (1987) und Li (2003) vor, dass die Oberflächenkontur des Neutraltons als Interpolation zwischen zwei Tontargets⁸ des vorausgehenden und des nachfolgenden Tons⁹ realisiert wird. Die jüngste Arbeit über den Neutralton von Chen & Xu (2006) zeigt, dass die variable Oberflächenkontur des Neutraltons von dem vorausgehenden Ton bestimmt wird. Darüber hinaus argumentieren die Autoren, dass der Neutralton ein grundlegend statisches Tontarget in der Mitte des Stimmumfangs zu sein scheint. Der Neutralton wird im Vergleich zu Silben mit vollen Tönen mit deutlich kürzerer Dauer und schwachem artikulatorischem Aufwand realisiert. Dabei ist die Dauer

⁸ Im Rahmen der Autosegmentalen Phonologie besteht die F0-Kontur einer Äußerung aus einer Reihe von tonalen Ereignissen (z.B. Tonakzente im Deutschen oder lexikalische Töne im SC). Die phonetischen Realisierungen solcher Ereignisse bezeichnet man als „Tontargets“.

⁹ Hier ist „Ton“ ein voller lexikalischer Ton (Shih 1987) oder ein postlexikalischer Ton, z.B. ein Grenzton (Li 2003).

ein ausschlaggebendes Merkmal für diesen Ton. Nach Yangs (2010) Studie hätte der Neutralton eine große Ähnlichkeit mit Ton 3 und Ton 4, wenn es keinen Dauerunterschied gäbe.

Resümee

Zusammengefasst kann man festhalten, dass die vier lexikalischen vollen Töne im SC zwei entscheidende Merkmale besitzen: das Register und die Kontur. Der Neutralton hat eine hoch variable Oberflächenkontur und wird mit weniger artikulatorischem Aufwand und kurzer Dauer realisiert.

2.1.2 Töne des Standardchinesischen im Kontext

Im obigen Abschnitt wurden alle Töne des SC detailliert in Zitierform beschrieben. Allerdings werden diese Töne im normalen Sprachgebrauch nur selten isoliert realisiert. Wenn lexikalische Töne im SC sowie in anderen ostasiatischen Sprachen in zusammenhängenden Äußerungen produziert werden, weichen die Tonhöhen und die Konturen der Töne beachtlich von ihren kanonischen Formen ab, ohne die Kontraste zwischen den Tönen zu verlieren. Dieses Phänomen wird normalerweise als tonale Koartikulation bezeichnet. Manche koartikulatorischen Abweichungen basieren auf Regeln oder sind mit bestimmenden Morphemen verbunden. Sie werden in der Forschung häufig als Tonsandhi bezeichnet. Im folgenden Abschnitt werden Tonsandhi-Phänomene und verschiedene tonale Variationen erläutert.

2.1.2.1 Tonsandhi

Tonsandhi beschreibt Tonveränderungen (nach Wiese 1988), die nach bestimmten Regeln und tonalem Kontext operieren. Perzeptiv wird der geänderte Ton von Muttersprachlern auch als eine andere tonale Kategorie wahrgenommen. Allerdings verwenden manche Autoren wie Yip (2002) und Chen (2000) den Terminus *Sandhi*, um weitere Phänomene zu beschreiben, wie etwa Allotöne oder

Tonveränderungen, die durch die Satzintonation und andere, z.B. morpho-syntaktische Bedingungen herbeigeführt werden.

T3-Sandhi

Die phonetische Realisierung des T3 weist viele Variationen auf. In Zitierform zeigen sogar Muttersprachler unterschiedliche Muster bei der Produktion von T3, wie in Abbildung 2.4 demonstriert:

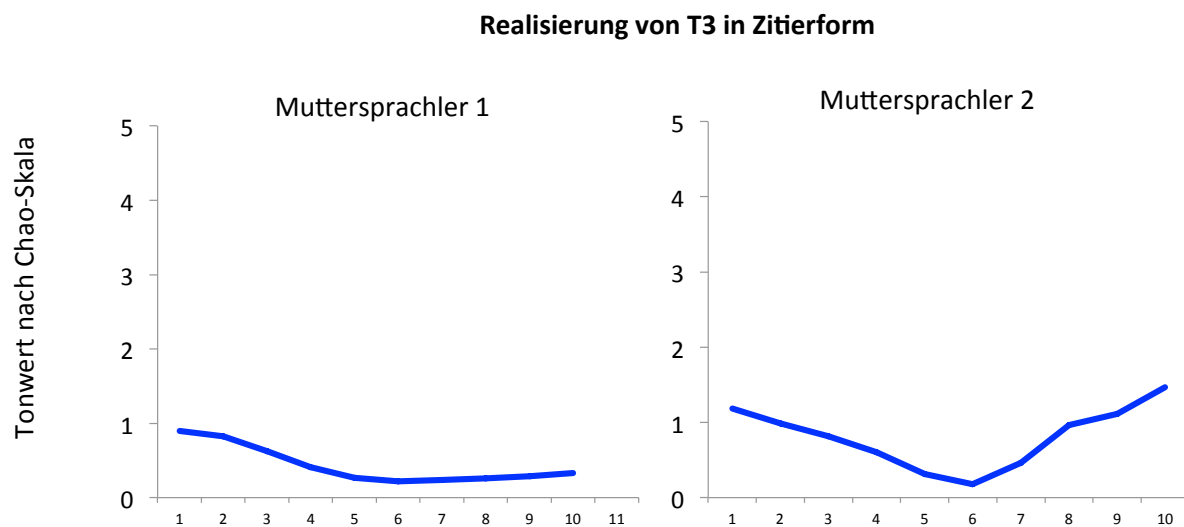


Abbildung 2.4: Unterschiedliche Realisierungen des T3 in Zitierform von zwei Muttersprachlern.

Die Zitierform des T3 in einer finalen Position (kein weiterer Ton folgt diesem T3) hat eine fallend-steigende Kontur und nach der Fünferskala von Chao den Wert /214/. In einer nicht finalen Position wird T3 meistens mit einer abweichenden Kontur realisiert. In diesem Fall wird T3 nur mit einem sogenannten „halben T3“ realisiert. Wenn T3 vor einem T1, T2, T4 oder T0 produziert wird, wird die steigende F0-Bewegung getilgt und hat den Wert /21/. Wenn T3 vor einem anderen T3 innerhalb eines prosodischen Fußes¹⁰ produziert wird, erfolgt die Realisierung des ersten T3 mit einer Kontur wie T2 /35/. Dieses Phänomen wird auch als T3-Sandhi bezeichnet (Chao 1968; Chen 2000; Duanmu 2000;), (2.8) zeigt ein Beispiel der T3-Sandhi.

¹⁰ In der Forschungsliteratur sind ausführliche Diskussionen über die Bedingungen und phonetischen Eigenschaften des T3-Sandhi zu finden (Chao 1968; Cheng 1973; Shih 1986; Kratochvil 1987; Zhang 1988; Hung 1989; Shen 1990b; Hsiao 1991; Chen 2000, unter anderen).

(2.8) Ton 3 + Ton 3 /214/ + /214/ → /35/ + /214/

- a. mǎi mǎi mǎ
 kaufen kaufen Pferd
 „kaufen“ „Pferd kaufen“
 /214/ /214/ /214/ Grundform
 /35/ /214/ Sandhi-Form
- b. mái mái mǎ
 begraben begraben Pferd
 „begraben“ „Pferd begraben“
 /35/ /35/ /214/ Grundform = Sandhi-Form

Variationen in der Realisierung von T3 werden in Tabelle 2.2 zusammengefasst:

	Allotöne		
	/214/	/21/	/35/
Sandhi	–	–	+
Finale Position	+	+	–
Anderenorts	–	+	–

Tabelle 2.2: Ton 3 Variationen nach Chen (2000: 20).

Da der Sandhi-T2¹¹ von Muttersprachlern perzeptiv meistens nicht von T2 in Zitierform differenziert werden kann (Peng 1996), wird diese Tonveränderung auch als „phonologisches“ Tonsandhi bezeichnet. Jedoch liefert Peng (1996) den Beweis für eine unvollständige Neutralisierung des T2 und Sandhi-T2, wie Abbildung 2.6 zeigt:

¹¹ Sandhi-T2 bezeichnet den durch die Sandhi-Regel geänderten T3, der von Muttersprachlern als T2 wahrgenommen wird.

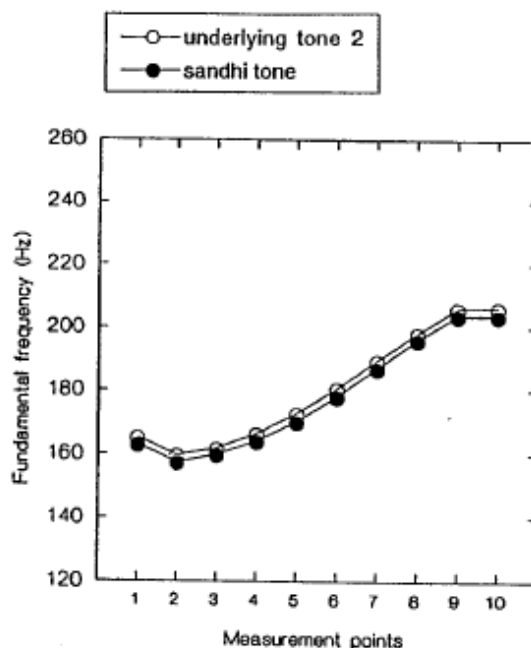


Abbildung 2.5: Phonetische Realisierung von T2 und Sandhi-T2 (Peng 1996: 22).

In Abbildung 2.5 kann man erkennen, dass die Konturen von T2 und Sandhi-T2 sehr nah beieinander liegen. Allerdings sind sie phonetisch nicht identisch und weisen ein Phänomen auf, das der „*incomplete neutralisation*“¹² der Auslautverhärtung im Deutschen ähnlich ist (Klerber 2011, Winter & Röttger 2011, unter anderen). Diese T3-Sandhi-Regel zählt zu den wenigen Regeln, die im ChaF-Unterricht expliziert erklärt und trainiert werden.

Andere Tonvariationen

Neben T3-Sandhi gibt es noch mehrere lokale Tonvariationen, wie etwa die sogenannte „morpho-phonemische“ Tonsandhi-Veränderung (Peng et al. 2005). Derartige Tonveränderungen kommen nur bei bestimmten Morphemen und tonalen Kontexten vor. Zum Beispiel wird der ursprüngliche T4 des Artikels *yì* („ein“) zu einem T2 vor einem anderen T4, *yì bēi* („eine Tasse ...“) vs. *yí kuài* („ein Stück ...“). Der T4 des Adverbs *bù* („nicht“) wechselt ebenfalls zu einem T2 vor einem anderen T4, *bùwǎn* („nicht spät“) vs. *búdùì* („nicht korrekt“).

¹² Der Begriff „*incomplete neutralisation*“ geht auf Port et al. (1981) zurück. Die Autoren bezeichnen den Prozess, der eine partielle Neutralisierung des Stimmhaftigkeitskontrastes in wortfinaler Position bewirkt, als *incomplete neutralisation* (Kleber 2011).

Es werden auch „phonetische“ Tonsandhi-Phänomene dokumentiert (Chao 1968; Duanmu 2000). Wie in Beispiel (2.9) gezeigt, wird ein T2 /35/ in einem dreisilbigen Wort zu einem T1 /55/, wenn T1 oder T2 vor diesem T2 stehen und ein beliebiger weiterer Ton nachfolgt.

(2.9)	rén	mín	bì	„Renminbi“ (Chinesische Währung)
	Mensch-Volk		Währung	
	/35/	/35/	/51/	Grundform
	↗	↗	↘	
	/35/	/55/	/51/	geänderte Form
	↗	—	↘	

Einige Studien haben versucht, die generellen Ursachen für Tonveränderungen dieser Art zu finden (Chen 2000; Shen 1992; Shih 1988; Wu 1982; Xu 1997). Ihre Ergebnisse deuten darauf hin, dass die tonale Koartikulation der wichtigste Faktor ist. Im nächsten Abschnitt wird deshalb die tonale Koartikulation detailliert erläutert.

2.2 Tonale Koartikulation

Viele Studien haben sich mit der tonalen Koartikulation in verschiedenen Sprachen auseinandergesetzt (Han & Kim 1974; Brunelle 2009; Abramson 1979; Gandour et al. 1992a, 1992b; Gandour et al. 1994; Peng 1997; Shih 1988; Shen 1990b; Xu 1994, 1997).

Die Studien zur tonalen Koartikulation haben sich hauptsächlich auf drei Aspekte konzentriert: die Richtung, den Wirkungsbereich und das Wesen der tonalen Koartikulation. Es herrscht allgemein Einigkeit darüber, dass zwei Richtungen der tonalen Koartikulation existieren. Der Einfluss eines unmittelbar vorausgehenden Tons wird in der Literatur als **progressiver** Effekt bezeichnet. Wenn der Einfluss jedoch von dem unmittelbar nachfolgenden Ton kommt, wird dies als **regressiver**

Effekt bezeichnet. Der progressive Effekt ist in vielen Tonsprachen dominant. Die Dominanz eines der beiden Effekte hängt jedoch von der jeweiligen Sprache ab. Der Wirkungsbereich der beiden Effekte ist nicht nur sprachspezifisch, sondern hängt auch von diversen Faktoren ab, wie dem Sprechtempo (DeCanio 2014), der Stimmhaftigkeit des zwischenstehenden Konsonanten in einem zweisilbigen Wort (Gandour 1992a, 1992b, 1994) oder dem Fokusstatus der betreffenden Konstituenten (Xu 1994). Das Wesen der tonalen Koartikulation variiert ebenfalls in verschiedenen Sprachen. Zumeist ist der progressive Effekt assimilatorisch und der regressive Effekt dissimilatorisch. Es gibt auch Sprachen (z.B. Taiwanesisch), in der umfangreiche regressive Assimilation zu finden ist (Peng 1997).

Da ein wichtiger Teil der vorliegenden Arbeit darauf abzielt, die tonale Koartikulation in der L2-Produktion zu erforschen, ist es notwendig, die vorherige Forschung zur tonalen Koartikulation der unterschiedlicher Tonsprachen zu besprechen und für die vorliegende Arbeit relevante Forschungsinhalte und Methoden darzustellen. Im folgenden Abschnitt werden einige Studien zur tonalen Koartikulation in ostasiatischen Tonsprachen vorgestellt.

2.2.1 Tonale Koartikulation in ostasiatischen Tonsprachen

Palmer (1969) beginnt seine Studie über Thailändisch mit Beobachtungen aus der Lehrpraxis. Die Lerner berichten demnach häufig, dass Ton A in unterschiedlichen Worten unterschiedlich klinge. Palmer findet heraus, dass die Realisierung des hohen Tons am Anfang eines zweisilbigen Wortes von der Identität des nachfolgenden Tons relativ unbeeinflusst ist, während die Realisierung desselben Tons am Ende eines zweisilbigen Wortes von dem vorausgehenden Ton erheblich beeinflusst wird. Der Effekt ist zum Teil so stark, dass sich die Richtung des Tonhöhenverlaufs ändert. Dieses Phänomen ist in der Tonproduktion des SC ebenfalls zu finden (Chao 1968; Xu 1997). Zudem behauptet Palmer, dass die Variation der Höhe und der Auslenkung des Tons von dem vorausgehenden und nachkommenden Ton abhängig ist. Obwohl der Umfang seiner Studie begrenzt

ist, zeigt er bei der Realisierung des hohen Tons einen stärkeren progressiven als regressiven Effekt auf.

Abramson (1979) stellt in seinen Daten zu den fünf lexikalischen Tönen und ihren 25 Kombinationen im Thailändischen schwache progressive und regressive Effekte fest. Die nicht so starken Effekte sind auf die nicht kontrollierten Konsonanten und Wortakzente der Komposita zurückzuführen, die in seiner Studie verwendet werden. In seinen Daten scheinen manche lexikalischen Töne wegen der metrischen Struktur neutralisiert zu sein, sodass einige tonale Kontraste verloren gehen. Diese Studie basiert allerdings nur auf visuellen Beobachtungen ohne eine quantitative statistische Analyse.

Die umfangreichste und ausführlichste Forschung zur tonalen Koartikulation im Thailändischen führt Gandour (1992a, 1992b und 1994) durch. In den beiden Studien im Jahr 1992 stellt er fest, dass die koartikulatorischen Effekte in zweisilbigen Wörtern im Thailändischen asymmetrisch sind. Der regressive Effekt ist stärker als der progressive. Im Vergleich zu früheren Studien setzt Gandour hier quantitative Messungen und eine inferentielle statistische Analyse ein. In einer Studie im Jahr 1994 findet Gandour jedoch einen stärkeren progressiven als regressiven Effekt. Die Unterschiede zwischen seinen Studien könnten durch das unterschiedliche Testmaterial verursacht worden sein. Die Stimmhaftigkeit des zwischenstehenden Konsonanten ist ein Hauptfaktor, der den Wirkungsbereich der tonalen Koartikulation mitbestimmt. Stimmlose Plosive scheinen die Koartikulation zu blockieren.

Die tonale Koartikulation im Vietnamesischen untersuchen Han & Kim (1974) und finden heraus, dass Höhe und Kontur der Töne außerordentlich von der tonalen Umgebung modifiziert werden. Die Variation eines Tons ist höher, wenn er neben Tönen mit hohen Onset-Werten steht. Die hochvariablen Konturen der Töne werden nicht überschritten, d.h. wenn der Onset-Wert eines Tons durch den Einfluss des vorangehenden hohen Tons erhöht ist, würde der Offset-Wert auch

erhöht und nicht reduziert. Im Vietnamesischen ist der progressive Effekt größer als der regressive.

Marc Brunelle (2003; 2009) forscht systematisch an Nord- und Südvietnamesisch. Nordvietnamesisch zeigt eine stärkere progressive als regressive Koartikulation. Der progressive Effekt ist stark, aber schnell nachlassend und ist eher von lokaler Natur. Der regressive Effekt ist zwar schwach, hat aber eine längerfristige Wirkung. Das heißt, die Auswirkung des regressiven Effektes reicht bis zum Onset der betroffenen Silbe.

Es gibt auch Sprachen, die andere koartikulatorische Muster aufweisen, wie zum Beispiel das malaysische Hokkien – ein südchinesischer Dialekt. Chang und Hsieh (2012) zeigt in einer Studie, dass der progressive Effekt nicht generell der dominanter ist. Der progressive Effekt ist nicht zwingend assimilatorisch und der regressive muss nicht automatisch dissimilatorisch sein. Das malaysische Hokkien hat ein komplexes Tonveränderungssystem. Töne können nur in einer finalen Position einer Domäne in ihren Grundformen realisiert werden. In allen anderen Positionen werden nur geänderte Töne realisiert. Dies wird als *Final Prominence* bezeichnet (Yip 2002). Die Symmetrie der progressiven und regressiven Koartikulation könnte aufgrund der *Final Prominence* in der Süd Min Sprache zustande gekommen sein. Die starke finale prosodische Position in dieser Sprache versichert einerseits, dass der Ton in dieser Position seine Grundform erreicht, verhindert aber andererseits den progressiven Einfluss des vorausgehenden Tons.

Die Studie von Peng (1997) über Taiwanesisch ergibt, dass F0-Variationen der Töne in dieser Sprache primär von der prosodischen Position abhängig sind. Taiwanesisch hat sieben lexikalischen Töne, und Tonsandhi ist prosodisch kontextsensitiv. Der Ton der vorletzten Silbe einer prosodischen Phrase wird systematisch zu einer Sandhi-Form geändert, sodass ein Abfall der F0-Kontur in der finalen Position der Phrase auffällig ist. Je höher die Position der betreffenden Silbe innerhalb der prosodischen Hierarchie ist, desto größer ist der koartikulatorische Effekt auf diese Silbe, sofern sie in einer finalen Position steht.

Das heißt, ein Effekt ist in einer äusserungsfinalen Position stärker als in einer phrasenfinalen. Peng stellt unter den Tönen im Taiwanesischen fest, dass die nicht fallenden Töne in allen tonalen Kontexten marginal stabiler sind, als die fallenden Töne. Am stabilsten sind die steigenden Töne.

2.2.2 Tonale Koartikulation im SC

In diesem Abschnitt wird tonale Koartikulation im SC detailliert erläutert. Frühere Studien von Shih (1987) und Wu (1984; 1988) haben festgestellt, dass die F0-Kontur eines bestimmten Tons in unterschiedlichen tonalen Umgebungen variiert. Die koartikulatorischen Effekte im SC beeinflussen die Onset- und Offset-Werte eines Tons. Zum Beispiel weist Wu (1988) darauf hin, dass der Offset-Wert des ersten T4 in einer T4-T4 Sequenz bei weitem nicht so tief fällt wie in der Zitierform.

Shen (1990) untersucht dreisilbige tonale Sequenzen im SC mit den Silben /pa pa pa/ in einem Trägersatz: *zhè shì ... sān gè zì* „Hier sind ... drei Charakter“. Die zu testende Silbe befindet sich in der Mitte des dreisilbigen Wortes. Shen misst den Onset-Wert, den Offset-Wert und den Wendepunkt des jeweiligen Tons. Als Ergebnis findet sie heraus, dass es im SC sowohl regressiven als auch progressiven koartikulatorischen Effekt gibt, die beide symmetrisch sind. Die koartikulatorischen Effekte im SC beeinflussen die gesamte F0-Kontur, jedoch nicht die Auslenkung des Tons. Die symmetrische Koartikulation in Shens Studie ist möglicherweise darauf zurückzuführen, dass die Nonsense-Silben /pa pa pa/ von den Sprechern sorgfältig und gleichmäßig betont produziert werden. Das heißt, die Betonung der Silben wird künstlich modifiziert. In einem natürlichen dreisilbigen Wort im SC wird jedoch die letzte Silbe am meisten und die mittlere Silbe am wenigsten betont.

Xu entdeckt in seiner Studie (1994) zu dreisilbigen Wörtern und Phrasen, deren mittlere Silbe entweder T2 oder T4 trägt, eine asymmetrische Verteilung zwischen progressiven und regressiven Effekten in der tonalen Koartikulation des SC. Xu unterteilt die tonalen Kombinationen der vier lexikalischen Töne im SC in zwei

Kontexte: kompatibel (*compatible*) und inkompatibel (*conflicting*). In einem kompatiblen Kontext haben die benachbarten Töne identische oder ähnliche phonetische F0-Werte, wie zum Beispiel den hohen Offset-Wert des T1 und den hohen Onset-Wert des T4 in einer T1-T4 Sequenz. In einem inkompatiblen Kontext haben die benachbarten Töne sehr unterschiedliche phonetische F0-Werte, zum Beispiel in einer T1-T2 Sequenz, in der der Offset-Wert von T1 hoch und der Onset-Wert von T2 niedrig ist. Die Zielsilben in Xus Studie stehen entweder in einem kompatiblen oder inkompatiblen Kontext. In einem kompatiblen Kontext sind die Auslenkungen (*slope*) der Konturtöne (T2 und T4) viel steiler als in einem inkompatiblen Kontext, in dem T2 und T4 flache Konturen haben, zum Teil zeigt T2 sogar eine fallende statt einer kanonisch steigende Kontur. Jedoch verursacht die radikale Abweichung der Kontur des T2 keine Verlagerung bei der tonalen Wahrnehmung, d.h. die fallende Kontur wird immer noch als T2 wahrgenommen. Generell weichen die Töne in einem inkompatiblen Kontext weiter von ihren Zitierformen ab als in einem kompatiblen Kontext. Die Auswirkung des tonalen koartikulatorischen Effekts wird laut Xus Studie größer, wenn sich das Sprechtempo erhöht.

In Xus Studie (1997) liegt der Schwerpunkt darin, das Wesen der progressiven bzw. regressiven tonalen Koartikulation zu identifizieren. In dieser Studie testet er deshalb zweisilbige Wörter (meistens Nonsense-Wörter) mit der Struktur /*ma ma*/ mit 16 möglichen Tonkombinationen in Trägersätzen. Bei der progressiven Koartikulation wird in den meisten Fällen ein assimilatorischer Effekt festgestellt und dieser ist in erster Linie auf den Offset-Wert des vorausgehenden Tons zurückzuführen. Das Ausmaß des assimilatorischen Effektes nimmt im Zeitverlauf ab. Im Gegensatz zur progressiven findet Xu nur geringfügige regressive tonale Koartikulation in seinen Daten, wobei diese dissimilatorisch zu sein scheint. Wenn ein beliebiger Ton vor Tönen mit niedrigem Onset-Wert (T2 und T3) steht, ist das F0-Maximum des Tons höher als wenn dieser Ton vor Tönen mit hohem Onset-Wert steht.

2.2.3 Ursprung und Einflussfaktoren der tonalen Koartikulation

In der wissenschaftlichen Diskussion gibt es unterschiedliche Erklärungsversuche für den Ursprung der tonalen Koartikulation, wobei viele Autoren in der Artikulation selbst den Ursprung sehen. Ein lexikalischer Ton kann nicht abrupt in den nächsten übergehen, sondern muss einen Übergang zwischen zwei Tönen haben. Die Tonproduktion basiert auf Variationen der Grundfrequenz, die wiederum hauptsächlich durch Muskelaktivitäten des Kehlkopfs erfolgt. Wenn ein Tontarget in ein anderes übergeht, benötigen die für die Phonation zuständigen laryngealen Muskeln ausreichend Zeit, um den Übergang zu vollziehen. Sprecher des SC brauchen durchschnittlich 142 ms zur Anhebung der Tonhöhe um sechs Halbtöne (Xu und Sun 2002). Die Trägheit der Muskeln könnte deshalb die Ursache für die progressive Assimilation in der tonalen Koartikulation sein. Die zweite mögliche Erklärung berücksichtigt die Vorausplanung der Sprachproduktion. Studien zeigen, dass die Koartikulation auf segmentaler Ebene meistens geplant ist (z.B. Whalen 1990). Bezüglich der tonalen Koartikulation behaupten Gandour et al. (1994), dass die regressive Dissimilation vorausgeplant sei, um den Übergang von einem hohen Tontarget zu einem sehr niedrigen zu erleichtern. Eine weitere Erklärung basiert auf der Perzeption. Peng (1997) und Xu (1994) haben bestätigt, dass Muttersprachler auf koartikulatorische Indizien zurückgreifen, um Töne zu identifizieren. Viele Autoren weisen darauf hin, dass die Motivation für die regressive Dissimilation eine Verstärkung der tonalen Kontraste bei der Perzeption ist (Xu 1997; Gandour et al. 1994; Potisuk 1997). Der exakte Ursprung der tonalen Koartikulation muss allerdings noch weiter erforscht werden.

Ebenso relevant wie die Erforschung des Ursprungs ist die Untersuchung möglicher Faktoren, welche die tonale Koartikulation der verschiedenen Sprachen beeinflussen. Von den inhärenten Unterschieden der Sprachen abgesehen, gibt es weitere Faktoren, welche die tonale Koartikulation beeinflussen. Ein Faktor ist die prosodische Position. Wenn ein T2 in einer nicht-betonten Position in dreisilbigen Wörtern/Phrasen steht, wird dieser oftmals mit einer inversen Kontur produziert

(Wu 1984 und Xu 1994). Dies widerspricht der Behauptung von Shen (1990), dass die Richtung eines Tons nicht von einem koartikulatorischen Effekt beeinflusst wird. Den Einfluss der prosodischen Position ist auch bei segmentaler Koartikulation zu finden (z.B. Cho 2004). Chen und Xu (2006) haben in einer Studie über den Neutralton gezeigt, dass sich der Einfluss eines vollen Tons nicht nur auf den nachfolgenden Neutralton beschränkt, sondern auch auf den zweiten oder dritten Neutralton ausdehnt. Dies widerspricht den Ergebnissen von Xu (1997) über SC und Potisuk et al. (1996) über Thailändisch, in denen kein tonaler koartikulatorischer Effekt über die benachbarte Silbe hinaus gefunden wird. Die unterschiedlichen Ergebnisse könnten auf die prosodische Stärke zurückzuführen sein, da in Xus und Potisuks Studien nur Silben an prosodisch starken Positionen untersucht werden. Andere Faktoren, wie das Sprechtempo (DeCanio, 2014) oder die Stimmhaftigkeit des zwischenstehenden Konsonanten in einem zweisilbigen Wort (Gandour 1992a, 1992b, 1994), beeinflussen tonale Koartikulation aber ebenfalls.

Der Effekt weiterer Faktoren, wie der Prominenz-Grad einer Konstituente für die tonale Koartikulation, ist bis jetzt wenig erforscht. Es wird allerdings berichtet, dass segmentale Koartikulation je nach Prominenz-Grad variiert, z.B. in einer Intonationsphrase akzentuiert oder nicht-akzentuiert (Cho 2002).

2.2.4 Verschiedene Modelle zu tonaler Koartikulation

Im Gegensatz zur tonalen wird die segmentale Koartikulation seit Jahrzehnten intensiv erforscht und zahlreiche Autoren haben verschiedene Modelle zu Koartikulation entwickelt. In den 1960er Jahren existieren Modelle, die sich hauptsächlich mit der temporalen Domäne der Koartikulation beschäftigen. Vertreter sind Kozhevnikov und Christovich (1965), die postulieren, dass die größte Einheit in der Sprachproduktion hinsichtlich der Artikulation ein Syntagma sei. Ein Syntagma enthält, nach den beiden Autoren, ungefähr sieben Silben und kann ein Satz oder Teil eines Satzes sein. Die artikulatorische Basiseinheit innerhalb eines Syntagmas ist die artikulatorische Silbe, die aus einem

Vokal und einer beliebigen Anzahl aufeinanderfolgender Konsonanten besteht. Koartikulation findet zwischen dem Ende des vorauslegenden Vokals und dem Start des nächsten Vokals statt, ungeachtet der Anzahl an Konsonanten, die dazwischen stehen. Es gibt andere Modelle, nach denen die Koartikulation sich auf die phonem-spezifischen Merkmale konzentriert, z.B. das „*coarticulation resistance*“ (CR) Konzept (Bladon und Al-Bamerni 1976) oder die „*feature based*“ Modelle (Keating 1990, unter anderen). In den 1990er Jahren hat Manuel das „*output constraints*“ Model vorgeschlagen. Sie untersucht Vokalsysteme und die phonetische Kontraste der Vokale verschiedener Sprachen und zeigt, dass Sprachen mit fünf phonemischen Vokalen stärkere Vokal-Koartikulation „tolerieren“ als Sprachen mit sieben phonemischen Vokalen. Die *output constraints* sind somit von den sprachspezifischen Kontrasten des phonetischen Systems beeinflusst.

Zusammengefasst lässt sich feststellen, dass die meisten Modelle lediglich Phänomene auf der segmentalen Ebene erklären. Brunelle (2009) und DiCanio (2014) verknüpfen ihre Daten mit dem *output constraints* Model von Manuel. Brunelle vergleicht die tonale Koartikulation im Nordvietnamesischen mit dem Südvietnamesischen und stellt einen stärkeren koartikulatorischen Effekt im Nordvietnamesischen fest. Während tonale Kontraste des Südvietnamesischen beinahe ausschließlich von der Grundfrequenz abhängen, spielt dieser akustische Parameter eine kleinere Rolle im tonalen Kontrast des Nordvietnamesischen, denn drei der sechs Töne im Nordvietnamesischen unterscheiden sich zusätzlich durch die Stimmqualität. Die Wahrscheinlichkeit, die Töne im Nordvietnamesischen aufgrund von Tonhöhe und Kontur zu verwechseln ist kleiner als im Südvietnamesischen. Dies führt dazu, dass Nordvietnamesisch eine ausgeprägtere tonale Koartikulation erlaubt. DiCanio (2014) geht noch einen Schritt weiter und behauptet, dass tonale Koartikulation vom Gewicht der einzelnen Komponenten des Tons abhängig sei. Tonale Koartikulation ist seiner Meinung nach in Tonsprachen eingeschränkter, in denen die Tonhöhe mehr Gewicht im tonalen Kontrast hat als in jenen, in denen die Tonkontur wichtiger ist.

Flemming (2011) geht ebenfalls von einem *constraint*-basierten Ansatz aus und analysiert tonale Koartikulation als ein Ergebnis der Interaktionen zweier Arten von *constraints*: artikulatorische und perzeptiv motivierte. Er stellt die These auf, dass tonale Koartikulation zum einen von universalen *constraints*, wie den physiologischen, und zum anderen von sprachspezifischen Tontargets bestimmt wird. Diese zwei Arten von *constraints* interagieren und verursachen die Vielfalt der Muster tonaler Koartikulation in verschiedenen Tonsprachen. Die physiologischen *constraints*, die auch mit artikulatorischen *constraints* gleichgesetzt werden, zeigen das einheitliche Phänomen, dass tonale Koartikulation in allen Tonsprachen existiert. Die sprachspezifischen *constraints* aufgrund unterschiedlicher Tontargets hängen äußerst eng mit der perzeptiven Identifikation der Töne zusammen. Abbildung 2.6 stellt diese zwei Arten von *constraints* schematisch dar, denen tonale Koartikulation unterliegt.

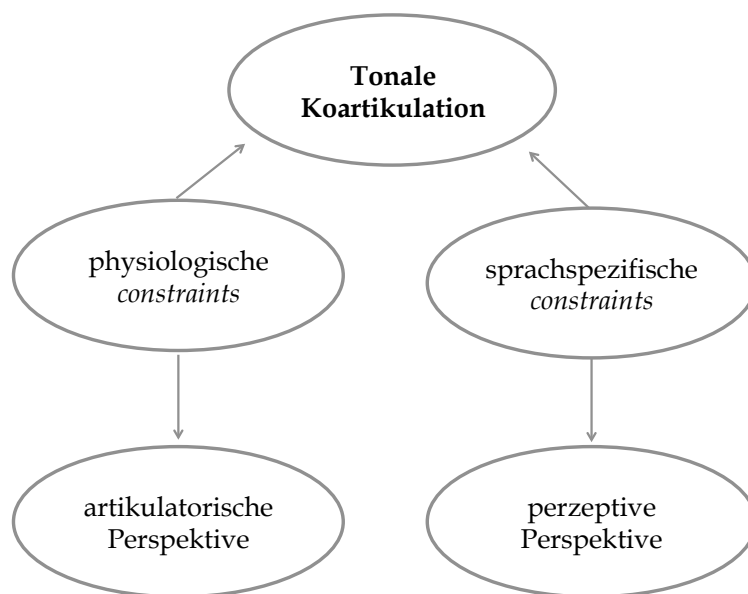


Abbildung 2.6: Darstellung des *constraint*-basierten Modells der tonalen Koartikulation nach Flemming (2011).

Wenn man die Theorie von Flemming für die vorliegende Studie adaptiert, müsste tonale Koartikulation in der L2-Produktion im SC ebenfalls den universalen *constraints* unterliegen, weil auch deutsche Lerner einen lexikalischen Ton nicht ohne einen tonalen Übergang realisieren können. Wenn unterschiedliche Muster oder Ausmaße der tonalen Koartikulation in der L1- und L2- Produktion zu

finden sind, ist das höchstwahrscheinlich auf unterschiedliche Reaktionen auf die sprachspezifischen *constraints* zurückzuführen.

Resümee

Tonale Koartikulation ist ein universales Phänomen in Tonsprachen. Die Richtung, der Wirkungsbereich und das Wesen tonaler Koartikulation sind aufgrund der unterschiedlichen Tontargets sprachspezifisch und hängen zusätzlich von verschiedenen Faktoren ab. Im SC ist der progressive Effekt der dominanteren, der hauptsächlich das Onset der betreffenden Silbe beeinflusst und assimulatorisch ist.

2.3 Weitere Aspekte der Prosodie im Standardchinesischen

2.3.1 Wortakzent

Ein Wortakzent (*engl: lexical stress*) ist auditiv in den meisten westgermanischen Sprachen schnell und leicht identifizierbar. Aus diesem Grund besteht wenig Diskussion darüber, ob es in Sprachen wie Englisch oder Deutsch Wortakzente gibt. Dagegen sind die Autoren sich nicht einig darüber, ob Wortakzente im SC existieren und falls dies so ist, wo der Wortakzent in zweisilbigen Wörtern liegt.

Doch warum ist es so schwierig, Wortakzente im SC zu identifizieren? Standardchinesisch ist eine Tonsprache, in welcher der Tonhöhenverlauf nicht wie im Englischen oder Deutschen als Indikator eines Wortakzents verwendet werden kann. Bis dato gibt es jedoch eine Übereinstimmung unter den Autoren, dass Silben mit einem Neutralton metrisch schwach¹³ sind. Somit erhält die erste Silbe in einem zweisilbigen Wort den Wortakzent. Viel umstrittener ist jedoch die Frage, ob es Unterschiede hinsichtlich des Wortakzents bei Wörtern ohne Neutralton gibt. Es existieren drei Theorien über Wortakzente im SC (Duanmu 2000):

¹³ Im Rahmen der *Metrical Phonology* (Hayes 1995, Halle & Idsardi 1995, Halle 1998, unter anderen) ist das Alternieren zwischen starkem und schwachen Takt eine fundamentale Eigenschaft des Wortakzents oder Rhythmus (Duanmu 2000:126). Jedes Alternieren konstruiert einen Fuß, im dem die Silbe mit einem schwachen Takt als metrisch schwach bezeichnet wird.

(1) **Kein-Wortakzent-Theorie** (*No-Stress Theory*):

SC Muttersprachlern fällt es schwer, einen Wortakzent in der Sprache zu erkennen. Es gibt einige Paare zweisilbiger Wörter, in denen Wortakzente distinktiv sind, z.B. dà yì vs. dà yì. Der Ton von „yì“ wird „neutralisiert“ und trägt keinen Ton. Aus diesem Grund argumentieren manche Autoren (Gao und Shi 1963), dass die oben genannten zwei Wörter nicht durch den Wortakzent sondern das Vorhandensein eines Tons unterschieden werden.

(2) **Final-Wortakzent-Theorie** (*Final-Stress Theory*):

Chao (1968) und Xu (1980) sind sich einig, dass im SC die stärkste Prominenz auf der finalen Silbe liegt. Lin, Yan, und Sun (1984) berichten ebenfalls, dass in isolierten zweisilbigen Wörtern die zweite Silbe möglicherweise aufgrund einer finalen Dehnung (*final lengthening*) eine längere Dauer hat. Dieses Ergebnis bezieht sich allerdings nur auf die Realisierung von isolierten zweisilbigen Wörtern. Andere Studien (Feng 1985, Wang & Wang 1993) zeigen, dass zweisilbige Wörter in einer Äußerung häufig ein anderes Muster des Wortakzents zeigen als in Isolation. In diesem Fall wird meistens die Anfangssilbe statt der finalen Silbe mit längerer Dauer produziert. Es wird deutlich, dass es keinen Beweis mehr für die Final-Wortakzent-Theorie gäbe, wenn der Effekt der finalen Dehnung wegfiel.

(3) **Kein-Kopf-Wortakzent** (*Nonhead Stress*):

Cinque (1993) und Duanmu (1990) argumentieren, dass in Komposita und Phrasen die „*nonhead*“ Komponente mehr Prominenz als die „*headed*“ Komponente hat. Syntaktische „*nonheads*“ müssen einen Wortakzent tragen. In einem Kompositum im SC, das aus zwei „*nonhead*“ Monosilben besteht, hat die erste Silbe mehr Prominenz.

Duanmu (2000) stellt die These auf, dass ein Wortakzent im SC durch zwei Verfahren bestimmt werden kann. Ein einfaches zweisilbiges Wort bildet einen Trochäus, z.B. im Wort *qīngtíng* „die Libelle“ besteht keine interne syntaktische

Struktur zwischen der ersten und zweiten Silbe. In diesem Fall erhält die erste Silbe geringfügig mehr Prominenz als die zweite. Dagegen unterliegt die Akzentzuweisung bei Phrasen dem *Nonhead Stress* Prinzip, z.B. trägt in der Nominalphrase *gāo shān* („hoher Berg“) der nicht-Kopf *gāo* („hoch“) den Wortakzent.

Außerdem entdeckt Duanmu (2004) in der phonetischen Realisierung Ähnlichkeiten zwischen einem Trochäus im Englischen und Wörtern mit einem Neutralton im SC. Hier ist in beiden Sprachen die erste Silbe im Wort die metrisch stärkere. Diese Silbe ist länger und hat keine Vokalreduktion und wird durch die Dauer und Intensität im Englischen und einen lexikalischen vollen Ton im SC prominenter realisiert. Die schwächere (zweite) Silbe ist kürzer und hat reduzierte Vokale. Allerdings ist der tonale Status grundsätzlich unterschiedlich. Im Englischen hat die schwächere Silbe keine mit ihr verankerten tonale Spezifikation, während im SC diese Silbe einen Neutralton trägt. Obwohl der Neutralton in manchen Studien nicht als Ton gilt, besitzt er trotz der schwachen Implementierung ein Tontarget.

2.3.2 Funktionalität der Grundfrequenz

In den vorherigen Abschnitten wird lexikalische Prosodie beschrieben. In diesem Abschnitt werden Intonation auf der Äußerungsebene und die Interaktion zwischen Ton und Intonation erläutert.

In Sprachen wie Deutsch oder Englisch sind die wichtigsten phonologischen Spezifikationen die Tonakzente (*pitch accents*). In Sprachen wie dem SC sind dies die lexikalischen Töne. In beiden Sprachfamilien kann die F0-Kontur als eine Reihe von Tönen oder phonologischen Spezifikationen beschrieben werden. Der wesentliche Unterschied zwischen Tonsprachen und Intonationssprachen liegt nicht in der An- oder Abwesenheit der lexikalischen Töne, sondern in der Funktionalität der phonologischen Spezifikationen der jeweiligen Sprache.

Die Beschreibung der Intonation auf der Äußerungsebene im SC ist aufgrund der multiplen Funktionalität der Grundfrequenz komplizierter als die einer Intonationssprache. Im SC ist die lokale F0-Modifizierung begrenzt und benötigt andere Mechanismen, um die wichtigsten Funktionen der Intonation zu erfüllen, beispielsweise zum Phrasieren oder Hervorheben einzelner Komponenten in Äußerungen.

2.3.2.1 Satzintonation

In der Forschung wird dokumentiert, dass z.B. der Unterschied zwischen einer Frage und einer Aussage durch das globale Expandieren oder Komprimieren der F0, die Modifizierung der letzten Silbe in einer Äußerung oder eine globale Kontur-Veränderung intonatorisch markiert werden kann.

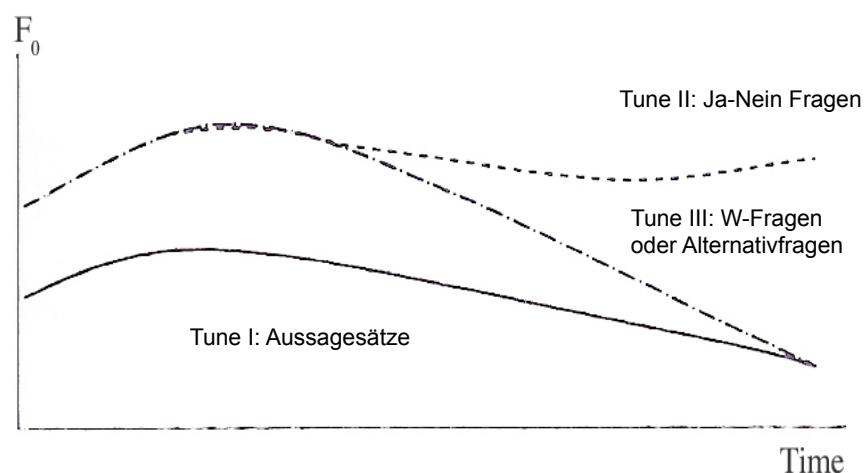


Abbildung 2.7: Globale Tonhöhenbewegungen in verschiedenen *tunes* nach Shen (1990a)

Shen (1990a) behauptet, dass es mindestens drei distinktive *tunes* globaler F0-Variationen gibt. Sie bezeichnet diese als *tune I* bei Aussagesätzen, *tune II* bei Ja-Nein Fragen und *tune III* bei W-Fragen oder Alternativfragen („Ist es A oder B?“), wie in Abbildung 2.7 illustriert. Dabei bestimmen drei Komponenten die jeweilige Form der *tunes*: Der Anfangspunkt (*starting point*), der Verlauf der F0-Kontur im mittleren Bereich des Satzes, und der Endpunkt (*ending point*). Tune II beschreibt sie beispielsweise wie folgt:

tune II: starting with a mid-high key, moving upward to a high key at the highest peak, dropping but not too low, ending in the high or mid-high register.

[Tune II: Startet mit einer mittelhohen Tonart, bewegt sich aufwärts zu einer hohen Tonart auf dem höchsten Gipfel, fällt, aber nicht so tief, endet (trotzdem) im hohen oder mittelhohen Register]

(Shen 1990a: 26, meine Übersetzung)

2.3.2.1 Interaktion zwischen Ton und Intonation

Lexikalische Töne als Grenztöne

Frühere Studien über die Interaktion zwischen Ton und Intonation bei verschiedenen Satzmodi können in zwei Thesen unterteilt werden. Die eine ist die *pitch-neutralization* These (Wu 1982, Shen 1990a unter anderen), die besagt, dass die Tonhöhe oder -kontur der lexikalischen Töne im SC von der Satzintonation modifiziert bzw. neutralisiert wird. Wenn T2 sich in einer fallenden Satzintonation befindet, wird dieser Ton mit weniger steiler Steigung realisiert, z.B. am Ende eines Aussagesatzes. Die andere ist die *pitch-succession* These (Chao 1968), die wiederum besagt, dass dem Tonhöhenverlauf eines betreffenden Grenztons eine zusätzliche tonale Bewegung hinzugefügt wird. Zum Beispiel wird ein T4 in einer steigenden Intonation mit einem Tonwert von /513/ statt /51/ realisiert, wodurch dieser Ton erheblich modifiziert wird.

Fokusmarkierung im Standardchinesischen

Eine Interaktion zwischen Ton und Intonation im SC wird in der Fokusmarkierung deutlich. Da die F0-Veränderungen im SC im Gegensatz zu westgermanischen Sprachen auf der lexikalischen Ebene bedeutungsunterscheidend sind, hat SC weniger Möglichkeiten, unterschiedliche pragmatische Funktionen durch F0-Manipulationen zu erfüllen. Seit langem wird in der Forschung berichtet, dass der Fokus im SC durch Modifikation des F0-Umfangs (*pitch range*) des betreffenden Tons markiert wird.

Gårding (1987) schlägt das Gittermodell (*range grid*) vor und stellt eine starke Asymmetrie zwischen den fokussierten und den post-fokalen Komponenten fest, nämlich dass die fokussierten Komponenten durch expandierten F0-Umfang manifestiert werden und die post-fokalen Komponenten eine beträchtlich post-fokale Kompression zeigen. Xu (1999, 2005) zeigt, dass die Bildung der F0-Kontur im SC von verschiedenen parallel wirkenden Faktoren bestimmt wird. Darunter führt die Fokusmarkierung nicht nur zur Modifikation des F0-Umfangs einzelner Komponenten, sondern beeinträchtigt auch die globale Kontur eines Satzes, wie in Abbildung 2.8 dargestellt:

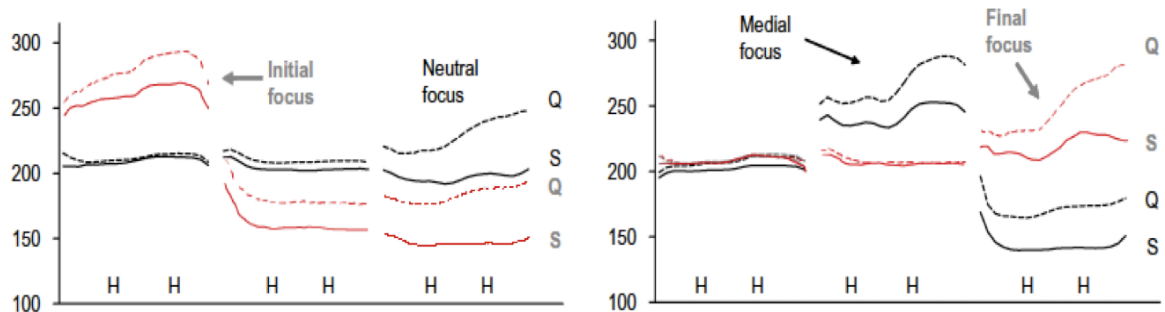


Abbildung 2.8: F0-Konturen von Aussagesätzen (S) und Fragen (Q) mit Fokus in verschiedenen Positionen: Am Satzanfang (initial focus), in der Mitte des Satzes (medial focus) und am Ende des Satzes (final focus), Xu (2005: 236).

Yiya Chen analysiert die Interaktion von Ton und Fokus gründlich (Chen 2003, 2010; Chen & Gussenhoven 2008; Chen & Braun 2006) und stellt die These auf, dass das Expandieren und Komprimieren des F0-Umfangs nicht als die primäre bzw. einzige Eigenschaft der tonalen Realisierung unter dem Einfluss der Fokusmarkierung betrachtet werden sollte. Zum Beispiel findet Chen (2012) heraus, dass ein T4 mit größerem F0-Umfang realisiert wird, wenn er in der post-fokalen Position hinter einem fokussierten T1 und vor einem T2 steht, als wenn er selbst im Fokus steht (siehe Abbildung 2.9).

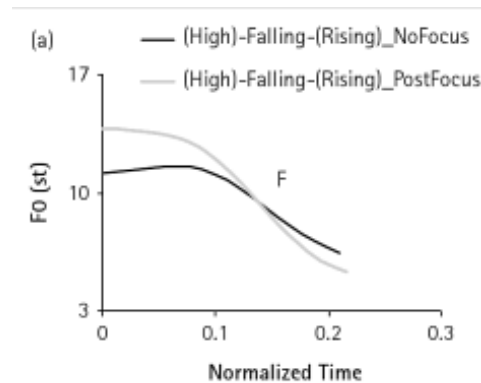


Abbildung 2.9: Beispiel für eine Realisierung des T4 mit einem größeren F0-Umfang in einer post-fokalen Position (Chen 2012: 112)

Chen & Gussenhoven (2008) stellen in ihren Daten fest, dass lexikalische Töne unter Emphase mit einer F0-Kontur realisiert werden, die den lexikalischen Kontrast verstärkt. Diese Hervorhebung wird den benachbarten Tönen und der verlängerten Silbendauer angepasst, sodass die maximal kontrastive Form eines Tons erreicht wird. Diese beiden Studien zeigen, dass die Realisierung eines Tons im Fokus von einer Reihe Faktoren als Ganzes beeinflusst wird, um den Kontrast dieses Tons zu maximieren. Die tonale Realisierung eines post-fokalen Tons wird als eine Manifestation der schwachen Implementierung eines Tons analysiert, weil dieser Ton eine prosodisch nicht-prominente Position einnimmt.

2.4 Deutsche Intonation

Intonation hat in Sprachen wie dem Deutschen, Englischen oder Niederländischen zwei Hauptaufgaben: (1) Hervorheben – das Prominenz-Verhältnis unter den Komponenten markieren – und (2) Phrasieren – den Redefluss in kleinere Einheiten unterteilen (Grice & Baumann 2007). In Abbildung 2.10 werden Prominenzen verschiedener Art schematisch dargestellt. Das Hervorheben einer bestimmten Silbe bei den oben genannten Sprachen wird entweder mit einer postlexikalischen Betonung (*stress*) oder einem Tonakzent (*pitch accent*) verschiedener Typen realisiert. Die lexikalische Betonung (Wortakzent) ist abstrakt und befindet sich auf der lexikalischen Ebene, auf den ein konkreter Tonakzent fallen kann. Wenn die betreffende Silbe mit einem Druckakzent

markiert ist, wird sie mit einer erhöhten Lautstärke und längeren Dauer sowie mit einem nicht-reduzierten Vokal produziert. Erhält diese Silbe zusätzlich eine F0-Bewegung, spricht man von einem Tonakzent.

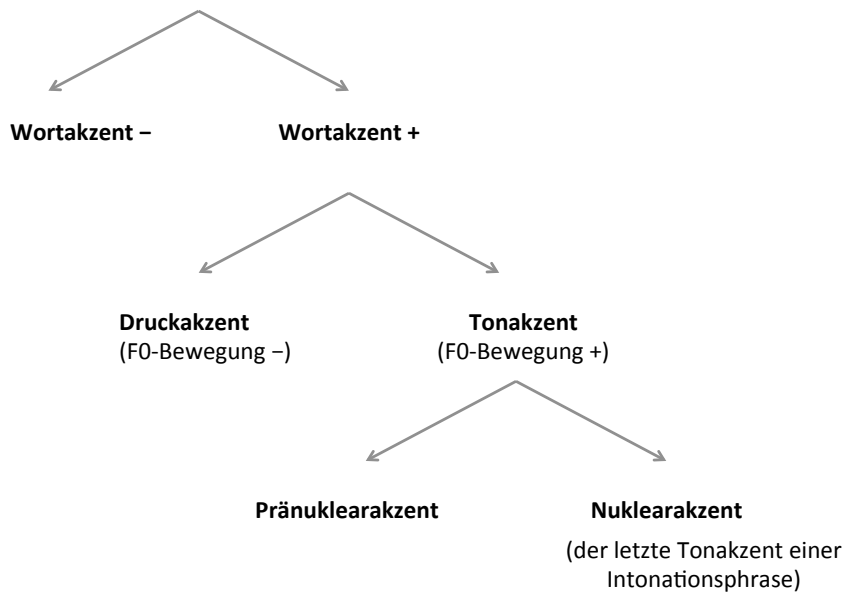


Abbildung 2.10: Schematische Darstellung unterschiedlicher Typen der Hervorhebung im Deutschen, adaptiert von Terken & Hermes (2000)

Zusammenfassend kann das Hervorheben in westgermanischen Sprachen in den folgenden Stufen erfolgen:

	Ebene	Prominenz-Status	Parameter
a. Wortakzent	Wortebene	abstrakt; kann eine konkrete Prominenz erhalten	_____
b. Druckakzent	Äußerungsebene	konkrete Prominenz	Dauer; Intensität
c. Tonakzent	Äußerungsebene	konkrete Prominenz	Dauer; Intensität; F0-Bewegung

Tabelle 2.3: Verschiedene Prominenzstufen (Grice & Baumann 2007)

Varianten (b) und (c) unterscheiden sich lediglich durch den Grad der (postlexikalischen) Prominenz und die F0-Bewegung.

2.4.1 Wortakzent

Es herrscht eine langjährige Debatte über die Zuweisung des Wortakzents im Deutschen. Hauptsächlich wird das Problem von zwei sehr unterschiedlichen Ansätzen analysiert: gewichtssensitive und gewichtsinensitive.

Die Vertreter des gewichtssensitiven Ansatzes sind zum Beispiel Giegerich (1985), Vennemann (1990) und Féry (1998). Demnach ist die Akzentzuweisung stark vom Silbengewicht abhängig. Hier wird zum einen zwischen langen und kurzen Vokalen unterschieden, dabei ist eine Silbe mit einem langen Vokal schwerer als eine mit einem kurzen. Zum anderen spielt die Silbenstruktur eine Rolle bei dem Silbengewicht. Silben mit Auslautkonsonanten, auch als geschlossene Silben bezeichnet, sind schwerer als Silben ohne Auslautkonsonanten, auch als offene Silben bezeichnet. Die Zuweisung des Wortakzents erfolgt auf Basis der Struktur der letzten oder vorletzten Silbe.

Giegerich schlägt eine *Main Stress Rule* vor:

$$(2.10) \quad \begin{array}{c} S \\ | \\ \text{syllable} \rightarrow \text{syllable} / ___ ((\text{light syllable})\text{light syllable}) \end{array}$$

Giegerich (1985:31)

Die Hauptakzentregel in (2.10) sagt aus, dass der am weitesten rechts stehende terminale Knoten „S“¹⁴ von einer oder zwei leichten Silben gefolgt werden kann, wie in den Worten „Genese“ in (2.11) b. und „Kompositum“ in (2.11) c. Fällt der „S“-

¹⁴ „S“ bedeutet *strong* (stark).

(2.11) **a.**

S W

De -kan Ø

b.

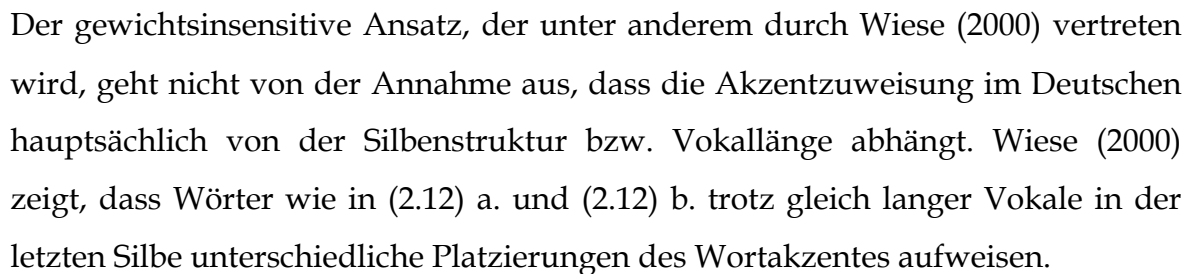
S W

Ge -ne se

c.

S
S W W

Kom -po si tum



- (Wiese 2000: 278)

(2.13)

a) **Fuß-Regel:** Füße des Typs F oder - wenn nicht möglich - F, werden

$$\begin{array}{ccc} & / \ \backslash & \\ s & w & s \end{array}$$

von rechts nach links gehend errichtet.

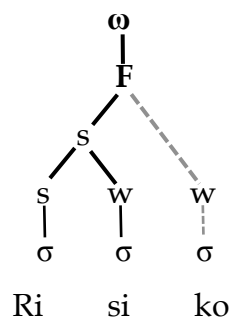
b) **Wort-Regel:** In einem phonologischen Wort ist der am weitesten rechts stehende Fuß stark.

- c) **Einbindungsregel:** Die verbliebenen Silben werden in minimaler Art einem Fuß als schwaches Mitglied eingebunden.

(adaptiert von Wiese 2000: 282)

Laut Wiese (2000) ist für die Akzentzuweisung im Deutschen entscheidend, einen zweisilbigen Fuß vom Typ *stark – schwach* aufzubauen und falls dies nicht möglich ist, einen einsilbigen Fuß vom Typ *stark* zu errichten. Den Wortakzent erhält der am weitesten rechts stehende Fuß, der stärker ist, als die vorangehenden Füße in einem phonologischen Wort. Extrametrische Silben, d.h. Silben, die nicht am Fußaufbau beteiligt sind, sollen in einen *strong-weak* Fuß eingebunden werden. Laut Regel (2.13) c. wird beispielsweise die finale Silbe „-ko“ in *Risiko* dem davorstehenden *stark – schwach* Fuß hinzugefügt, da sie eine extrametrische Silbe ist. Die interne Fußstruktur des Wortes *Risiko* demonstriert Beispiel (2.14).

(2.14)



Der einzige Zusammenhang zwischen Akzentzuweisung und Vokalqualität ist laut Wiese (2000), dass eine Silbe mit Schwa nie den Wortakzent bekommen kann, weil diese keinen Nukleus enthält.

Trotz stark voneinander abweichender Theorien zur Akzentzuweisung im Deutschen herrscht weitgehend Einigkeit darüber, dass das moderne Standarddeutsch eine trochäische Sprache ist und in der Regel die vorletzte Silbe den Wortakzent erhält, wenn die letzte Silbe reduziert ist (Giegerich 1995; Vennemann 1990; Janssen 2003 unter anderen).

2.4.2 Tonakzent

2.4.2.1 Akzenttypen

Das von Martine Grice, Matthias Reyelt, Ralf Benzmüller, Anton Batliner und Jörg Mayer gemeinsam erarbeitete Transkriptionssystem für die Intonation des Deutschen - GToBI (*German Tone and Break Indices*) wurde im theoretischen Rahmen der autosegmental-metrischen Phonologie entwickelt und ist für die Beschreibung des Intonationssystems des Deutschen weit verbreitet (Grice et al. 1996; Grice et al. 2005; Reyelt et al. 1996). Das GToBI-System enthält sechs elementare Tonakzente und elf nukleare Konturen, die aus einem Nuklearakzent und einem Grenzton zusammengesetzt sind. Silben mit Tonakzenten können unterschiedliche Tonhöhenbewegungen enthalten (Siehe Abbildung 2.11). Obwohl die „H“- bzw. „L“- Töne nicht als absolute F0-Werte sondern in Relation zum jeweiligen Stimmumfang eines Sprechers zu interpretieren sind, liegen H-Töne in der Regel etwa in den oberen drei Viertel und L-Töne im unteren Viertel des Stimmumfanges (Grice & Baumann 2002).

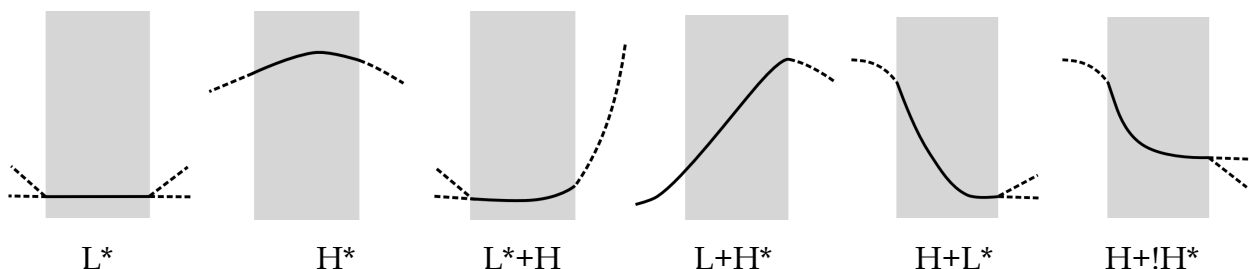


Abbildung 2.11: Sechs Tonakzente im Deutschen in GToBI (www.gtobi.uni-koeln.de)

Die Tonakzente sind entweder monotonal wie L* oder bitonal wie L*+H. Die grau markierten Bereiche in Abbildung 2.11 repräsentieren die akzentuierten Silben, mit denen ein gesternter Ton verankert ist. Die ausführlichen Beschreibungen der Tonakzente sind in Grice & Baumann (2002) oder der GToBI-Webseite (www.gtobi.uni-koeln.de) zu entnehmen. In Abbildung 2.12 und 2.13 werden zwei Tonakzente anhand ihrer F0-Konturen als Beispiele illustriert:

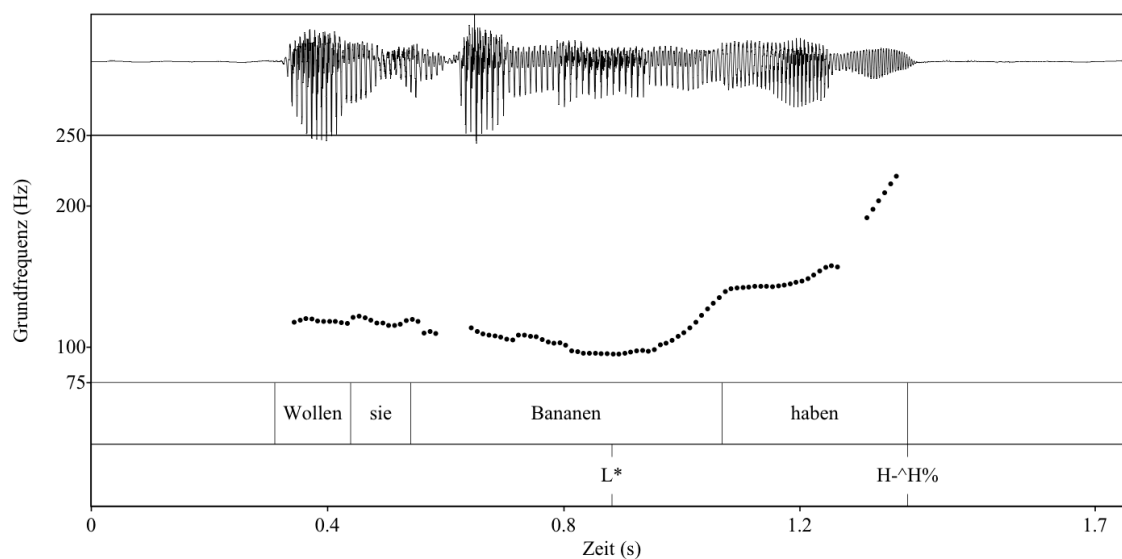


Abbildung 2.12: GToBI-Annotation zum Tonakzent L* (www.gtobi.uni-koeln.de)

In Abbildung 2.12 ist der Tonakzent L* mit der zweiten Silbe in „BaNAnen“ verankert. Die Silbe, die mit L* assoziiert ist, repräsentiert ein lokales F0-Minimum im Stimmumfang eines Sprechers. Dieser Tonakzent kann einer leicht fallenden Tonbewegung folgen, wie die auf der Silbe „Ba-“ in *Bananen*. Bei diesem Akzenttyp sind neben der F0-Bewegung zusätzlich die akustischen Parameter Intensität und Dauer für die Wahrnehmung ausschlaggebend.

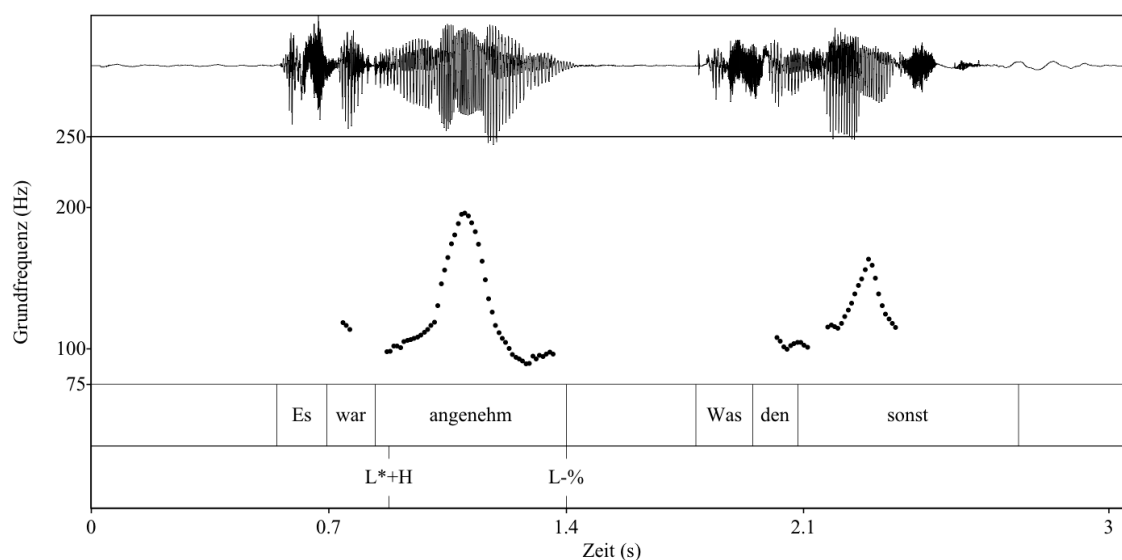


Abbildung 2.13: GToBI-Annotation zum Tonakzent L*+H (www.gtobi.uni-koeln.de)

In Abbildung 2.13 ist der bitonale Tonakzent L*+H mit der ersten Silbe in „ANgenehm“ verankert. Im Vergleich zum Tonakzent L* folgt dem tiefen Ton in L*+H ein steiler Anstieg, der spät in der Silbe „AN-“ beginnt und seinen Gipfel in der nachfolgenden Silbe „-ge-“ erreicht.

2.4.2.2 Funktionen

In der gesprochenen Sprache erfüllt die Intonation diverse linguistische und paralinguistische Funktionen. Grice und Baumann (2007) fassen die Funktionen und ihre intonatorischen Realisierungen in der folgenden Abbildung 2.14 zusammen.

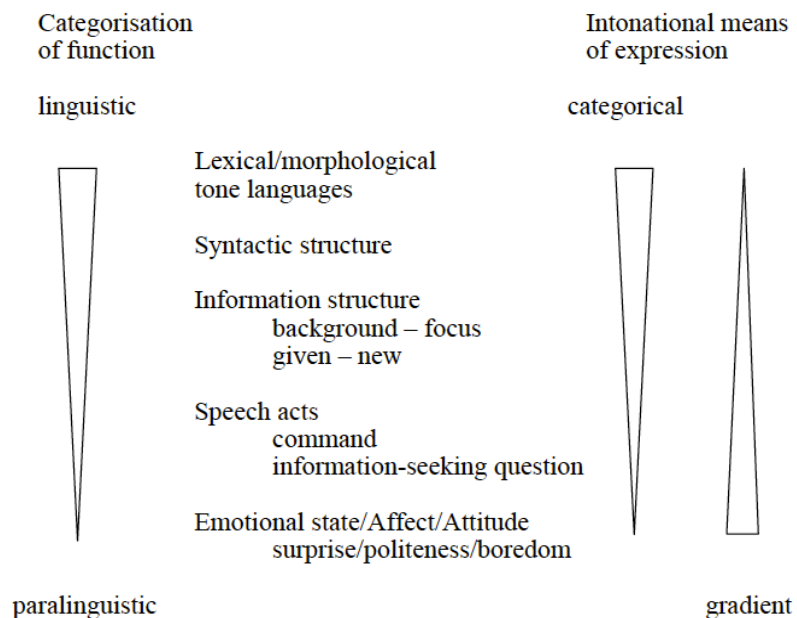


Abbildung 2.14: Prosodische Funktionen und ihre intonatorischen Realisierungen.
Grice & Baumann (2007)

Die Intonation in Sprachen wie Deutsch und Englisch kann Zweideutigkeiten mancher syntaktischer Strukturen auflösen, wie in Beispiel (2.15):

(2.15) Rainer verfolgt den Mann mit dem Motorrad.

(a). [Rainer]_{NP} [verfolgt]_{VP} [den Mann mit dem Motorrad]_{NP}

(b). [Rainer]_{NP} [verfolgt den Mann]_{VP} [mit dem Motorrad]_{PP}

(Grice & Baumann 2007)

Der Satz in (2.15) kann zwei Lesarten haben: (a) Rainer verfolgt den Mann, der mit einem Motorrad fährt. (b) Rainer verfolgt den Mann, indem Rainer selbst ein Motorrad fährt. Durch unterschiedliche Phrasierung des Satzes kann die Zweideutigkeit der syntaktischen Struktur beseitigt werden.

Die Intonation kann auch die Informationsstruktur markieren, z.B. den Grad der Gegebenheit, wie in Beispiel (2.16) gezeigt:

- (2.16) (a). Thomas hat Hunger. Also isst er einen **Apfel**.
 (b). A: Hast Du gesagt, dass Thomas mit einem Apfel jongliert?
 B: Nein, er **ISST** einen Apfel.

(Grice & Baumann 2007)

Während „Apfel“ in (2.16) a. einen Tonakzent erhält, ist dieses Wort in der Antwort in (2.16) b. deakzentuiert. Deakzentuierung beschreibt die Abwesenheit eines Tonakzentes an einer Stelle, wo dieser erwartet worden wäre (Baumann 2006). Das Verhältnis zwischen neuen und gegebenen Informationen wird dadurch deutlich.

Zusammengefasst liegt der entscheidende Unterschied zwischen SC und Deutsch hinsichtlich der Form der F0-Kontur in der Elastizität. Im SC ist eine bestimmte Kontur innerhalb einer Silbe zu realisieren. Im Deutschen ist ein bestimmtes Tontarget mit einer betonten Silbe assoziiert und die Kontur zwischen zwei Targets verläuft wie eine Interpolationslinie. Die F0-Kontur eines lexikalischen Tons hat im SC eine ausschließlich linguistische Funktion, während sie im Deutschen linguistische und paralinguistische Funktion erfüllen kann, wie Ladd (2008) dies anhand der Sprachen Chinesisch und Englisch verdeutlicht:

More „happens“ in pitch contours in Chinese, because the lexical tones occur at nearly every syllable and the transition between them span only milliseconds, whereas in English the pitch accents occur mostly only on prominent words, and the transition may span several syllables. But the specifications are the same kind of

phonological entity regardless of their function, and the transitions are the same kind of phonetics phenomenon irrespective their length.

Ladd (2008:148)

[Es „geschieht“ mehr in F0-Konturen im Chinesischen, weil die lexikalischen Töne beinahe in jeder Silbe auftreten und der Übergang zwischen ihnen nur Millisekunden dauert, während Tonakzente im Englischen meist nur auf prominenten Wörtern vorkommen und sich der Übergang über mehrere Silben erstrecken kann. Jedoch sind die Spezifikationen ungeachtet ihrer Funktion von derselben Art phonologischer Einheiten, und die Übergänge sind unabhängig von ihrer Länge dieselben phonetischen Phänomene.]

(Ladd 2008: 148, meine Übersetzung)

Ein Vergleich von Deutsch und SC wird in Tabelle 2.4 dargestellt:

		Deutsch	SC
Tonale Spezifikationen	Kategorien	Tonakzente	Lexikalische Töne
	Akustische Parameter	F0	F0
	Assoziation	betonte Silben	jede Silbe ¹⁵
	Implementierungsdomänen	eine bis mehrere Silben	eine Silbe
	Funktionen	linguistisch und paralinguistisch	linguistisch
Wortakzent in zweisilbigen Wörtern	Existenz	ja	ja/nein
	Akustische Parameter	Dauer; Intensität	Dauer; Intensität

Tabelle 2.4: Vergleich der prosodischen Systeme des SC und des Deutschen

¹⁵ Hier werden Silben mit Neutralton nicht berücksichtigt, da der Neutralton keine phonologische Spezifikation besitzt.

3 Theorien und Modelle im Zweitspracherwerb

In diesem Kapitel werden für die vorliegende Arbeit wichtige theoretische Grundlagen im Bereich Fremdspracherwerb vorgestellt, wie die *Interlanguage Hypothesis* oder die Theorie des Transfers sowie Einflussfaktoren im Ausspracheerwerb. Anschließend werden die Unterrichtspraxis und aktuelle Lehrmaterialien im Bereich ChaF beschrieben. Einige ausgewählte Studien zum Tonerwerb im SC werden danach vorgestellt. Dabei wird die Notwendigkeit der Untersuchung grundlegender Mechanismen der Lerner bei der Tonproduktion aufgezeigt.

3.1 Interlanguage - Lernersprache

Der Begriff *Interlanguage* - zu Deutsch „Lernersprache“ (Kniffka & Siebert-Ott 2007: 44) - ist auf Selinker (1972) zurückzuführen. Die Lernersprache beschreibt „das sich entwickelnde L2-System eines Lerners“ (Kniffka & Siebert-Ott 2007: 44). Obwohl ein Erwachsener beim Zweit- oder Fremdspracherwerb häufig Äußerungen mit fehlerhafter Grammatik oder starkem Fremdakzent produziert, sind derartige Äußerungen dennoch repräsentativ für ein zugrundeliegendes Sprachsystem. Major (2001) postuliert, dass die Lernersprache eine Kombination aus drei Bestandteilen besteht: L1, L2 und Universalien, die aus dem generellen Spracherwerbsprozess entstehen.

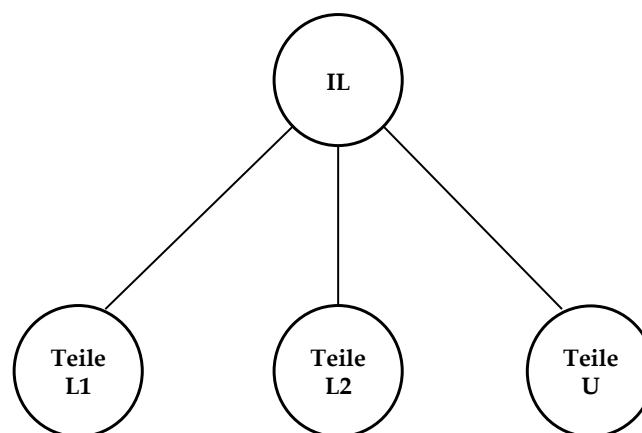


Abbildung 3.1: Schematische Darstellung der Bestandteile einer Lernersprache (Major 2001: 6)

Die nicht-zielsprachlichen Charakteristika der Lernersprache sind Forschungsgegenstand vieler L2-Studien. Fehlerhafte Äußerungen sind häufig auf Einflüsse der L1 zurückzuführen. Dies wird als „negativer Transfer“ (Major 2001) bezeichnet. Grundsätzlich kann Transfer auf allen linguistischen Ebenen vorkommen, beispielsweise in Form von Übertragungen muttersprachlicher Begriffe oder einer in der L1 üblichen Wortstellung in die L2.

Wenn ein Phänomen in L1 und L2 unterschiedlich ausgeprägt ist, geschieht häufig negativer Transfer und das Ergebnis sind Fehler bzw. Abweichungen von der Norm des L2-Systems. Im nächsten Abschnitt wird der negative Transfer genauer erläutert. Die Lernersprache kann ebenfalls Elemente der L2 enthalten.

Sofern Elemente aus L1 und L2 identisch sind, müssen Lerner diese nicht erneut erlernen. Dies bezeichnet man in der Zweit -oder Fremdsprachenforschung als „positiven Transfer“ (Major 2001).

Zusätzlich gibt es in einer Lernersprache Elemente, die weder mit negativem Transfer zu erklären, noch Elemente aus L2 sind. Zum Beispiel kann ein chinesischer Englisch-Lerner den Satz „*Does he gose to School?*“ produzieren. Da keine Flexion im Chinesischen vorkommt, kann dieser Fehler nicht durch den Einfluss der L1 entstanden sein. Es ist auch keine wohlgeformte Äußerung in der Zielsprache. Ein solches Phänomen könnte eine Folge der Universalien des Spracherwerbsprozesses sein. Wenn Lerner mit unterschiedlichem sprachlichen Hintergrund die gleichen Fehler beim Erwerb derselben Struktur der L2 machen, wie Kinder beim Erwerb der Zielsprache als Muttersprache, kann man annehmen, dass derartige Fehler eine Folge der sprachlichen Universalien sind, d.h. mit dem Prozess der Spracherwerbs an sich zu tun haben (Major 2001).

Bezüglich der Lernersprache ist hervorzuheben, dass diese ein individuelles linguistisches System ist. Es ist kein Zielsprachensystem, wie Englisch oder Deutsch, das man erlernen kann, sondern eine sich immer weiter entwickelnde Sprache eines jeden Individuums. Die individuelle Variation in der Lernersprache

kann durch den Einfluss verschiedener Faktoren entstehen, z.B. das Alter der Lerner beim Beginn des Spracherwerbs, die Motivation oder die Kompetenz in der Muttersprache.

3.2 Negativer Transfer

Transfer kann auf jeder linguistischen Ebene vorkommen. Somit kann ein negativer Transfer bei der sprachlichen Produktion der Lerner auf jeder Ebene zu Fehlern führen. Die zentrale Idee der Forschung zum Transfer basiert auf einem kontrastiven Vergleich der betroffenen Sprachen.

Transfer kann durch Ähnlichkeiten zwischen Sprachen erklärt werden. Im *Perceptual Assimilation Model (PAM)* von Best und Kollegen (Best 1993 und 1994; Best & Tyler 2007) behaupten sie, dass die in der frühen Kindheit etablierten linguistischen Kategorien später eine nicht muttersprachliche Kategorie aufgrund der Ähnlichkeit zu ihr assimilieren, sodass ein L2-Phonem bei der Perzeption einem auditiv ähnlichen L1-Phonem zugeordnet und als dieses wahrgenommen wird. Diesen Prozess bezeichnet Kuhl (1992; Kuhl & Iverson 1995) als den muttersprachlichen Magnet-Effekt (*Native Language Magnet Effect*).

In Fleges (1995, 1999) *speech learning model (SLM)* wird postuliert, dass die Laute, die in der L1 wenig kontrastiv jedoch in L2 Phoneme sind, schwierig adäquat in L2 wahrgenommen werden können. Zum Beispiel sind für englische Muttersprachler der stimmhafte alveolare Plosiv /d/ und der stimmhafte dentale Frikativ /ð/ kontrastiv, was für spanische Muttersprachler nicht gilt. Für einen spanischen Englisch-Lerner ist es daher besonders schwierig, diese zwei Laute in L2 auseinanderzuhalten. Allerdings behaupten einige Autoren, dass nicht allein die Ähnlichkeiten die Schwierigkeiten beim Erwerb der L2 vorhersagen können. Beispielsweise beschreiben Bohn und Flege (1992), dass manche deutschen Englisch-Lerner ihrer L1 ähnliche Laute sogar besser in der Zielsprache produzieren können.

Die Vorhersage potentieller Schwierigkeiten oder negativen Transfers benötigt außer einer Ähnlichkeitsanalyse noch eine Analyse der Strukturunterschiede der betreffenden Sprachen. Die *Markedness Differential Hypothesis* von Eckman (1977) beschreibt die Markiertheit als Grundprinzip, um die Schwierigkeiten beim L2-Erwerb vorherzusagen und zu erklären, warum manche Strukturen aus der L1 transferiert werden und andere nicht. Markiertheit bedeutet für Eckman folgendes:

„... a phenomenon A in some language is more marked than B if the presence of A in a language implies the presence of B; but the presence of B does not imply the presence of A.”

[...ein Phänomen A in einer bestimmten Sprache ist markierter als B, wenn die Präsenz von A in einer Sprache die Präsenz von B impliziert; jedoch die Präsenz von B die Präsenz von A nicht impliziert.]

(Eckman 1977: 320, meine Übersetzung)

Die lexikalischen Töne sind beispielsweise markierter als die Intonation, weil das Vorhandensein der lexikalischen Töne das Vorhandensein der Intonation impliziert, aber nicht umgekehrt. Die Markiertheit kann sich auch auf die Frequenz mancher Laute beziehen, z.B. ist /r/ markierter als /l/ im Amerikanischen Englisch.

Die MDH sagt bezüglich des Spracherwerbs voraus, dass:

a) Those areas of the target language which differ from the native language and are more marked than the native language will be difficult.

[Die Bereiche in der Zielsprache, die sich von der Muttersprache unterscheiden und markierter sind als in der Muttersprache, werden schwierig sein.]

- b) *The relative degree of difficulty of the areas of the target language which are more marked than the native language will correspond to the relative degree of markedness.*

[Der relative Schwierigkeitsgrad von den Bereichen der Zielsprache, die markierter sind als in der Muttersprache, entsprechen dem relativen Grad der Markiertheit.]

- c) *Those areas of the target language which are different from the native language, but are not more marked than the native language will not be difficult.*

[Jene Bereiche der Zielsprache, die sich von der Muttersprache unterscheiden, aber nicht markierter sind als in der Muttersprache, werden nicht schwierig sein.]

(Eckman, 1977: 321, meine Übersetzung)

Während zum Beispiel ein deutscher Englisch-Lerner Schwierigkeit hat, einen stimmhaften Obstruenten im Silbenauslaut zu produzieren, hat ein englischer Deutsch-Lerner damit keine Probleme.

Zusammengefasst beschäftigen sich die oben genannten Modelle überwiegend mit segmentalen Transferphänomenen. Jedoch bieten sie für die vorliegende Arbeit eine theoretische Basis, um potentielle Transferphänomene auf der supra-segmentalen Ebene zu analysieren.

3.3 Phonologischer und Phonetischer Transfer

Mennen (2007) appelliert für eine differenzierte Behandlung der phonologischen und phonetischen Komponenten in der Intonation einer Sprache, weil verschiedene Sprachen sich auf beiden Ebenen unterscheiden können. Phonologische Unterschiede stammen aus Unterschieden im Inventar der

phonologischen Einheiten, ihren Formen und den damit verbundenen Bedeutungen.

Cross-linguistic intonational differences at the phonological level are differences in the inventory of phonologically distinct tunes or categorical elements, their form, or in the meanings assigned to them (Ladd 1996).

[Sprachübergreifende intonatorische Unterschiede auf der phonologischen Ebene sind Unterschiede im Inventar von phonologisch distinktiven Tonbewegungen oder kategorischen Elementen, ihrer Form, oder in der ihnen zugeschriebenen Bedeutung (Ladd 1996).]

(Mennen 2004: 545, meine Übersetzung).

Deutsch und SC verwenden beispielsweise unterschiedliche phonologische Einheiten, um einen engen Fokus zu markieren. Während ein enger Fokus im Deutschen unter anderem durch die Anwesenheit der Akzenttöne markiert wird, ist die Modifikation des Stimmumfangs einer der Indikatoren für die Fokusmarkierung im SC.

Phonetische Unterschiede können aus Unterschieden der phonetischen Realisierung einer in L1 und L2 identischen phonologischen Einheit herrühren. Mennen (2004) untersucht die unterschiedliche phonetische Realisierung der phonologisch steigenden Kontur im modernen Griechisch und im Niederländischen. In beiden Sprachen kommt die gleiche Kontur vor und enthält identische phonologische Komponenten. Die zeitliche Alignierung des Gipfels erfolgt im Niederländischen allerdings früher als im Griechischen. Zusätzlich wird die Alignierung des Gipfels im Niederländischen von der Vokallänge der akzentuierten Silbe beeinflusst, während dies im Griechischen nicht der Fall ist.

Wenn es Unterschiede zwischen zwei Sprachen auf phonologischer und phonetischer Ebene gibt, kann Transfer aus der L1 auf beiden Ebenen erfolgen. In

der Aussprachedidaktik ist von großer Bedeutung, die Quelle der Einflüsse zu unterscheiden. Dadurch können Sprachlehrende mögliche Fehlerquellen besser und präziser identifizieren und daraufhin die Fehler gezielter behandeln.

3.4 Einflussfaktoren beim Ausspracherwerb

Bisher wurden sprachspezifische Faktoren erläutert, welche die Aussprache einer Lernaltersprache beeinflussen. In diesem Abschnitt werden weitere Faktoren aufgezeigt. Faktoren wie die Lernsituation, das Alter des ersten Sprachkontakts, die Dauer des Aufenthaltes im Zielsprachland, die Motivation und andere haben nachweislich eine Auswirkung auf den Erwerb der L2-Aussprache (Flege et al. 1995; Flege & Fletcher 1992, Missaglia 1999; Bongaerts et al. 1997, unter anderen).

3.4.1 Lernsituation: Zweitsprache vs. Fremdsprache

Beim L2-Spracherwerb unterscheidet man zwischen Zweitsprache und Fremdsprache, doch häufig wird beides gleichermaßen als L2 bezeichnet. Anhand mehrerer Kriterien kann man jedoch eine Unterscheidung treffen. Hierzu gehören z.B. die Kommunikationsfunktion der Zielsprache, der Lernort und der Grad der Steuerung des Lernens durch institutionelle Richtlinien (Dahmen 2013).

Nach Edmondson und House (2006: 9) handelt es sich bei der L2 um eine Zweitsprache, wenn sie die Kommunikationssprache ist, die für das alltägliche Leben eine unverzichtbare Rolle spielt. Dagegen nehmen Henrici und Vollmer et al. (2001: 8) die Unterscheidung durch den Lernort vor. Wenn die L2 innerhalb der Zielkultur erworben wird, spricht man demnach von Zweitspracherwerb. Falls der L2-Erwerb in der Ausgangskultur stattfindet, spricht man von Fremdspracherwerb.

Der Grad der Steuerung durch institutionelles Lernen ist auch ein wichtiger Parameter. Häufig ist der Fremdspracherwerb ein institutionell gesteuertes

Lernen, das heißt, das Lernen findet üblicherweise in einer Institution statt (Schule, Sprachschule, Sprachlernzentrum etc.). Der Lerner nimmt am Unterricht teil und hat in der Regel außerhalb des Unterrichtsraums wenig bis keinen Kontakt zu Muttersprachlern. Der Zweitspracherwerb hat dagegen immer eine natürliche Komponente (Rösler 1994: 7). Hierbei ist häufig der Anteil des institutionell gesteuerten Lernens gering. Der Sprachinput und -output in der zielsprachlichen Umgebung ist somit höher.

Allerdings erfolgt institutionell gesteuertes Lernen und natürlicher, ungesteuerter Erwerb häufig parallel. Zum Beispiel können Fremdsprachkurse in der Zielkultur im Rahmen von Austauschprogrammen oder Sprachreisen stattfinden. Umgekehrt könnte ein Einwanderer im Zielsprachland Sprachunterricht an einer Institution erhalten.

3.4.2 Alter des ersten Sprachkontakts

Über das Alter des ersten Sprachkontakts wird viel in der Forschung diskutiert. Lang herrschte die *Critical Period Hypothesis* (CPH, dt: Annahme der kritischen Periode). Diese Hypothese geht davon aus, dass eine Person einer Sprache innerhalb eines bestimmten kritischen Zeitfensters ausgesetzt sein muss, um diese auf muttersprachlichem Niveau erlernen zu können. Erfolgt dies erst nach dieser kritischen Periode, ist eine muttersprachliche Kompetenz nicht mehr zu erreichen (z.B. Lenneberg 1967; Scovel 1969, 1988). Es wird jedoch viel debattiert, ob eine kritische Periode tatsächlich existiert und bis zu welchem Alter sie dann reicht. Zudem gibt es unterschiedliche CPs (*Critical Periods*) bezüglich unterschiedlicher sprachlicher Fähigkeiten. Im Vergleich zu erwachsenen Lernern ist eine akzentfreie L2-Ausprache nach Long (1990) noch erreichbar, wenn der Beginn des Lernens vor dem sechsten Lebensjahr liegt.

3.4.3 Dauer des Aufenthaltes im Zielsprachland

Die Dauer des Aufenthaltes im Zielsprachland wird in der Literatur als *length of residence* (LOR) bezeichnet. Die LOR beschreibt die Jahre oder Monate, die ein Lerner in einer sprachlichen Gemeinschaft verbringt, in der die L2 als Sprache dominiert (Piske et al. 2001). Flege (1988) behauptet, dass die sprachliche Kompetenz bei den Lernern am Anfang ihrer Ankunft in der Zielkultur eine bemerkenswerte Verbesserung zeigt - ein Effekt, der dann relativ schnell abnimmt. Die LOR spielt keine große Rolle bei der L2-Aussprache eines erwachsenen Lerners. Später stellen Flege und Fletcher (1992) die These auf, dass die LOR einen geringeren Effekt im Vergleich zum AOL (*age of learning*) beim Erwerb der L2-Aussprache hat.

Die oben aufgelisteten Faktoren werden für alle Probanden in dieser Arbeit über ein Formular abgefragt. Bei der Auswahl der Probanden wird auf größtmögliche Homogenität geachtet und alle Informationen zu den Probanden werden im Kapitel 4 zusammengefasst.

3.5 ChaF im Unterricht

Die lexikalischen Töne des SC stellen selbst nach jahrelangem Lernen eine große Herausforderung für die Lerner dar, deren L1 keine Tonsprache ist (White 1981). Das liegt nicht nur an der Tatsache, dass die Töne des SC für westeuropäische Lerner tatsächlich eine Schwierigkeit darstellen, sondern auch an einer häufig ungeeigneten Art der Vermittlung im Unterricht. Die Schwierigkeit für englischsprachige Lerner besteht zum größten Teil in einer ungewohnt schnellen Änderung der Grundfrequenz bei jeder Silbe (Yang 2011).

Die Thematik der Töne wird oftmals lediglich einmalig zu Beginn des Kurses behandelt. Im chinesischen Studium vieler Hochschulen in Deutschland werden Phonologie und Phonetik des SC beispielsweise nur in den ersten Wochen unterrichtet. Im späteren Verlauf des Studiums werden Aussprachübungen nur

noch ab und zu durchgeführt. Solche Übungen setzen jedoch ausschließlich auf die Methode „Hör zu und sprich nach“.

Die Realisierung des Neutraltons wird im ChaF-Unterricht in der Regel nicht explizit trainiert. Den Lernern wird beigebracht, dass Silben mit einem Neutralton ohne eine tonale Spezifikation und kürzer produziert werden sollen. Die Lerner erhalten jedoch kein weiteres Training. Diese Instruktion kann die Lerner in sofern verwirren, dass nahezu jeder westeuropäische Lerner davon ausgeht, dass SC eine Tonsprache ist und jeder einzelnen Silbe ein lexikalischer Ton zugeschrieben wird. Wenn sie aufgefordert werden, eine Silbe ohne einen lexikalischen Ton zu produzieren, fehlt ihnen jedoch eine geeignete Strategie.

Hinsichtlich der Vermittlung der Töne im ChaF-Unterricht stellen sich zwei Fragen, die Jane Orton im Einführungstext zum Buch „*Research in Chinese as a Second Language*“ ausführlich beschreibt und diskutiert (Orton 2013). Zum einen fragen viele Lerner und Lehrer: Sind schlecht realisierte Töne von Bedeutung? Angesichts der Tatsache, dass in China verschiedene Dialekte mit unterschiedlichen Tonsystemen gesprochen werden, sehen viele Lerner keinen Nutzen darin, die Töne des SC perfekt zu beherrschen. Manche ChaF-Lehrer halten es sogar für ausgeschlossen, dass westeuropäische Lerner die lexikalischen Töne des SC überhaupt erwerben können (Orton 2013:20). Hinsichtlich des Spracherwerbs ist die Richtigkeit der Töne jedoch von großer Bedeutung. Dafür gibt es mehrere Gründe: (1) Für Lerner ist es essentiell, die Töne korrekt zu produzieren, da es die große Anzahl an Homophonen im SC für alle Sprecher notwendig macht, die Töne auseinanderzuhalten. Wenn die Lerner die Töne nicht korrekt beherrschen, haben sie erhebliche Probleme mit der Verständigung. (2) Das Vernachlässigen der Töne erschwert es Lernern, einen umfassenden Wortschatz aufzubauen, da die Töne bedeutungsunterscheidend sind. (3) Das Nicht-Beherrschen der Töne beeinflusst die Lernmotivation negativ, was häufig ignoriert wird. Wenn ChaF-Lerner die Töne nicht adäquat produzieren können, fühlen sie sich erfolglos und werden daher zunehmend unmotiviert. Oftmals ist dies der Grund für Resignation und Beendigung des Lernens (Orton 2013: 21).

Die zweite Frage lautet, ob es Möglichkeiten gibt, die aktuellen Unterrichtsmodelle hinsichtlich der Phonetik im ChaF zu verbessern. Auch diese Frage kann bejaht werden. Viele Studien liefern Beweise dafür, dass expliziter Unterricht einen positiven Effekt auf die Aussprachekompetenz bei Fremdsprachenlernern haben kann (z.B. Cenoz & Lecumberri 1996; Bongaerts et al. 1997; Elliott 1997; Derwing et al. 1997; Derwing & Rossiter 2003).

Speziell für den Erwerb von Tönen im SC haben bereits die Studien von Wang, Jongman & Sereno (2003) und He & Wayland (2010) gezeigt, dass explizites Aussprachetraining eine deutlich erkennbare Wirkung bei den Lernern erzielen kann. Perzeptionstraining verbesserte nicht nur die korrekte Wahrnehmung der Töne, sondern auch ihre Produktion und umgekehrt.

Die oben genannten Autoren befassen sich nicht mit speziell angefertigtem Trainingsmaterial, sondern mit der allgemeinen Spracherfahrung im SC. Dies ermutigt viele Autoren im Bereich ChaF, weitere Basisforschung zu betreiben, damit spezielles Lehrmaterial entwickelt werden kann.

3.6 Lehrbuchmangel

Eng verbunden mit den unzureichenden Inhalten im ChaF-Unterricht ist ein auffallender Mangel an (qualifizierten) Lehrbüchern im Bereich ChaF. In diesem Abschnitt wird eine Auswahl an Lehrbüchern vorgestellt, die in deutschen Hochschulen verwendet werden.

In „*Das neue praktische Chinesisch 1*“ (2007) werden phonetische Übungen bis zur sechsten Lektion behandelt. In der ersten Lektion wird mit einer kurzen Beschreibung der Konsonanten und Töne begonnen. Die Beschreibung der Töne beschränkt sich allerdings jeweils auf die kanonische Form. Es gibt keine Abbildungen, um beispielsweise die Konturen und die relativen Positionen der vier Töne innerhalb des Stimmumfangs darzustellen. In den folgenden Lektionen

werden nach und nach einige wenige Tonsandhi-Regeln vorgestellt. Bezüglich des Neutraltons wird in diesem Lehrbuch kurz beschrieben:

„In der modernen chinesischen Sprache gibt es einige Silben, die nicht akzentuiert sind und tonlos ausgesprochen werden. Silben im Neutralton tragen kein Tonzeichen.“

(Das Neue Praktische Chinesisch 1: Lehrbuch 2007: 20)

In der Vorstellung wird dem Lerner keine Anleitung zur phonetischen Realisierung des Neutraltons gegeben.

In „*Hanyu Jiaocheng*“ (2006) wird die Phonetik des SC innerhalb der ersten fünf Lektionen vermittelt. Es beginnt in der ersten Lektion mit einer Abbildung der vier lexikalischen Töne und einer Tabelle aller möglichen Silben ohne lexikalische Töne im SC. Jedoch fehlt hier eine Anleitung für die tonale Produktion in Tonsequenzen. Phonetische Übungen werden nach dem Schema „*Hör zu und sprich nach*“ durchgeführt. Über den Neutralton wird berichtet, dass er kurz und unbetont ausgesprochen wird und die F0-Werte der Töne zusätzlich von den vorangehenden Tönen beachtlich beeinflusst sind. Allerdings wird den Lernern keine konkrete Beschreibung der Oberflächenkontur des Neutraltons dargeboten, wenn er nach einem der vier vollen Töne produziert wird.

In „*Chinesisch Erleben*“ (2005) wird die Phonetik des SC inklusive einer Abbildung der Töne und einer kurzen Beschreibungen der üblichen Tonsandhi-Regel und des Neutraltons zusammengefasst in der Einleitung vorgestellt, ohne jegliche konkrete Anweisung zur phonetischen Realisierung. Nach dieser Einleitung werden keine weiteren phonetischen Übungen eingebaut. Ein Vorteil dieses Buches besteht in den enthaltenen Hörübungen, die sich in jeder Lektion befinden. Obwohl diese Übungen nicht speziell für die Phonetik des SC entwickelt sind, zeigen sie die Wichtigkeit des Hörens, was in vielen anderen Lehrbüchern vernachlässigt wird.

In „*Chinesisch für Deutsche 1 – Hochchinesisch für Anfänger in Kurzzeichen*“ (2004) entfällt die Phonetik vollständig. Dieses Lehrbuch konzentriert sich ausführlich und gründlich auf die Grammatik des SC.

Das spezielle Lehrbuch für das Chinesisch-Aussprachetraining „*Lóng – Aussprachetraining*“ (2009) wird als Zusatzmaterial zu einem Hauptlehrbuch verwendet. In diesem Buch werden die Töne zuerst zusammen vorgestellt. Anschließend werden zuerst T1 und T4 trainiert. Das Training beinhaltet T1 und T4 in Zitiertform und den Tonsequenzen T1-T1, T1-T4, T4-T4 sowie T4-T1. Sukzessiv werden T3 und T2 hinzugefügt. Es fehlen jedoch explizite Anleitungen zur koartikulatorischen Änderung der Töne.

Resümee

Zusammengefasst werden im ChaF-Unterricht hauptsächlich isolierte lexikalische Töne gelehrt. Es gibt weder spezielle Übungen für Wörter mit einem Neutralton noch für koartikulierte Töne. Tonale Übungen auf höheren tontragenden Domänen, wie zum Beispiel in Phrasen oder Sätzen, finden nicht statt.

In den Lehrbüchern wird im Rahmen der Einführung der chinesischen Aussprache meistens das gleiche Schema angewandt: eine Liste von Segmenten im Dialekt von Beijing; eine Tabelle von fast 400 potentiellen Silben ohne lexikalische Töne; eine Abbildung, in der die Onset-Werte und Verläufe der vier Töne gezeigt werden; die drei am häufigsten auftretenden Tonsandhi-Regeln (T3-Sandhi und zwei Tonänderungsregeln von „不“ *bù*, und „一“ *yì*). Kontrastive Übungen zu Tönen in Sequenzen werden dagegen meistens nicht angeboten.

3.7 Studien über Chinesisch als Fremdsprache (ChaF)

Es existieren bereits einige Studien über den Erwerb der Aussprache im SC. In diesem Abschnitt werden die Methodik und Ergebnisse ausgewählter Studien zusammengetragen und analysiert.

White (1981) stellt in seiner impressionistischen Studie fest, dass die unterschiedlichen tontragenden Domänen der Hauptunterschied zwischen den Tonakzenten im Englischen und Tönen im SC sind. Die Implementierungsdomäne der lexikalischen Töne im SC beschränkt sich auf die Silbenebene, während tonale Spezifikationen im Englischen auf einer oder mehreren Silben realisiert werden können. Anhand einer Fehleranalyse findet White heraus, dass die größte Anzahl an Fehlern direkt auf die unterschiedlichen Wortbetonungsmuster im Englischen und SC zurückzuführen ist. In der phonetischen Realisierungen der englischen Wörter wird die Wortbetonung meistens mit einer Tonhöhenbewegung assoziiert. Aus diesem Grund nehmen englische Lerner Silben mit einem hohen Ton im SC als betonte Silben wahr. Dagegen verhalten sich unbetonte Silben im Englischen sehr ähnlich wie Silben mit einem T3 im SC. Deshalb werden Silben mit T3 im SC von englischen Lernern als unbetonte Silben wahrgenommen. Folglich zeigt sich häufig bei der Produktion von Lernern eine auffällige Vokalreduktion bei Silben mit T3. Allerdings basiert diese Studie auf einer theoretischen, kontrastiven Analyse des SC und des Englischen sowie auf Beobachtungen, ohne experimentelle Daten und statistische Analysen.

Miracle (1989) stellt fest, dass (US-)amerikanische Sprecher bei der Produktion zweisilbiger chinesischer Wörter mehr Fehler in der ersten als in der zweiten Silbe produzieren. Zehn Lerner und drei Muttersprachler nahmen an der Datenerhebung der Studie teil. Einsilbige und zweisilbige Wörter wurden aufgenommen, allerdings in der Fehleranalyse zusammengeführt, sodass separate Auswertungen zu Tönen in isolierter und koartikulierter Form nicht möglich sind. In der Fehleranalyse der Produktion der Töne unterteilt Miracle die Fehlertypen in zwei Kategorien: Registerfehler und Konturfehler. Daraus folgt der Befund in Tabelle 3.1:

	Registerfehler	Konturfehler	Verteilung der Fehlertypen
T1	F0 wird zu niedrig realisiert	durch eine fallende Kontur ersetzt	gleichmäßig verteilt
T2	F0 wird zu hoch realisiert	durch eine fallende und flache Kontur ersetzt	gleichmäßig verteilt
T3	F0 wird ausnahmslos zu hoch realisiert	durch eine abfallende Kontur ersetzt	gleichmäßig verteilt
T4	F0 wird zu niedrig realisiert	durch eine flache Kontur ersetzt	gleichmäßig verteilt

Tabelle 3.1: Zusammenfassung der Fehleranalyse nach Miracle (1989)

Zusätzlich entdeckt Miracle signifikant mehr Fehler in der ersten von zwei Silben bei zweisilbigen Wörtern. Allerdings stammen die drei teilnehmenden Muttersprachler vom chinesischen Festland und aus Taiwan. Taiwan-Chinesisch und SC weisen im tonalen Bereich eindeutige Unterschiede auf, was die Aussagekraft der Ergebnisse mindert. Zusätzlich werden die Testwörter an unterschiedlichen Positionen in die Trägersätze eingebaut, weshalb postlexikalische Einflüsse auf die tonale Produktion nicht kontrolliert werden können.

Shen (1989) zeigt in ihrer Studie, dass die (US-)amerikanischen Lerner am häufigsten Fehler bei T1 und T4 produzieren. Acht amerikanische Studierende des SC auf Anfängerniveau lesen eine ihnen vertraute Lektion aus dem Lehrbuch vor. Trotz fehlenden Vergleichs mit einer Kontrollgruppe von Muttersprachlern gibt Shen an, dass die Probanden T1 und T4 mit zu niedrigen F0-Werten produzieren. Daraus schließt sie, dass T1 und T4 für amerikanische Lerner die am schwierigsten zu erlernenden Töne sind. Die Fehlerquoten bei T2, T3 und T0 sind annähernd gleich. Jedoch haben die Probanden zur Zeit der Studie erst vier Monate SC gelernt und dies ist vermutlich der Grund, aus dem die Ergebnisse mit den meisten anderen Studien nicht übereinstimmen. In der Hauptsache wird berichtet, dass T2 und T3 Lernern die größten Schwierigkeiten bereiten (Chen 2000; Yang 2011; Hao 2012; unter anderen). Nichtsdestotrotz ist daran festzuhalten, dass die

Lerner in Anfängergruppen besonders darauf trainiert werden sollten, das richtige Tonregister zu erreichen.

Sun (1998) untersucht in ihrer Dissertation, wie (US-)amerikanische Lerner lexikalische Töne in einsilbigen und mehrsilbigen chinesischen Wörtern realisieren. Wie in vielen anderen Studien (Kiriloff 1969; Elliot 1991; Leather 1990; Chen 1997) erstellt Sun eine Rangordnung der relativen Schwierigkeiten im Zusammenhang mit den vier lexikalischen Tönen im SC. Das Ergebnis ihrer Studie zeigt, dass T2 und T3 die fehleranfälligsten Töne sind und deshalb als letzte erworben werden. Zusätzlich berichtet Sun, dass die Unterscheidung zwischen hohem und tiefem Ton die Hauptherausforderung für Lerner im Anfangsstadium darstellt, was die Befunde von Shen (1989) stützt. Die Probanden in Suns Studie produzieren einsilbige Wörter und die letzten Silben in mehrsilbigen Wörtern eindeutig mit höherer Genauigkeit als Silben in anderen Positionen. Sowohl in der Produktion als auch in der Perzeption ist die Verwechslung von T2 und T3 auffällig. Das bedeutet, dass die Konturen der beiden Töne verschmelzen, wenn die Probanden diese produzieren.

Die in der obigen Auswahl vorgestellten Studien sind auf die tonale Produktion auf der Silben- oder Wortebene fokussiert. Einige andere Studien analysieren tonale Produktion von Lernern auf der Äußerungsebene. In Chens Dissertation (2000), die sich auf (US-)amerikanische Chinesisch-Lerner bezieht, werden natürlich produzierte Äußerungen der Lerner verwendet. Bestimmte Tonsequenzen werden aus einer spontanen Äußerung extrahiert und von Muttersprachlern beurteilt. Bei diesem Versuchsaufbau ist zu erwarten, dass die Quote der als richtig empfundenen tonalen Kategorien bei den Muttersprachlern niedrig ist. Xu (1994) beweist, dass Muttersprachler des SC Informationen der benachbarten Töne benötigen, um Töne im Redefluss richtig zu identifizieren. Werden ihnen jedoch Töne präsentiert, die aus einer natürlich produzierten Äußerung ausgeschnitten sind, sinkt die Identifikationsquote. Basierend auf dem Urteil der Muttersprachler, beschreibt Chen die typische starke Verschmelzung von T2 und T3 sowie T4 und T1, wenn diese Tonpaare in einer Äußerung

produziert werden. Die erhobenen Sprachdaten zeigen die Tendenz, Konturtöne durch ebene Töne zu ersetzen. Bezüglich der Position neigt die Silbe in einer nicht finalen Position zu einem zu hohen und in einer finalen Position zu einem zu niedrigen F0-Wert.

Yang (2011) untersucht ebenfalls die tonale Produktion auf der Äußerungsebene durch (US-)amerikanischen Lerner und findet dabei einige allgemeine Fehlermuster. T2 und T3 sind für die Lerner die am schwierigsten zu produzierenden Töne und führen am häufigsten zu einer Fehlproduktion. Außerdem wird T2 am häufigsten in Äußerungsanfang oder -mitte und T3 häufig in der Äußerungsmitte produziert, wobei letzterer in der Regel als eine Pufferzone unmittelbar vor einem hohen Tontarget dient. Yangs Interpretation der Ergebnisse ist ein Transfer aus der englischen Intonation. Daraufhin schlägt er ein vorläufiges Modell des Tonerwerbs vor:

A tentative model of learners' tone errors

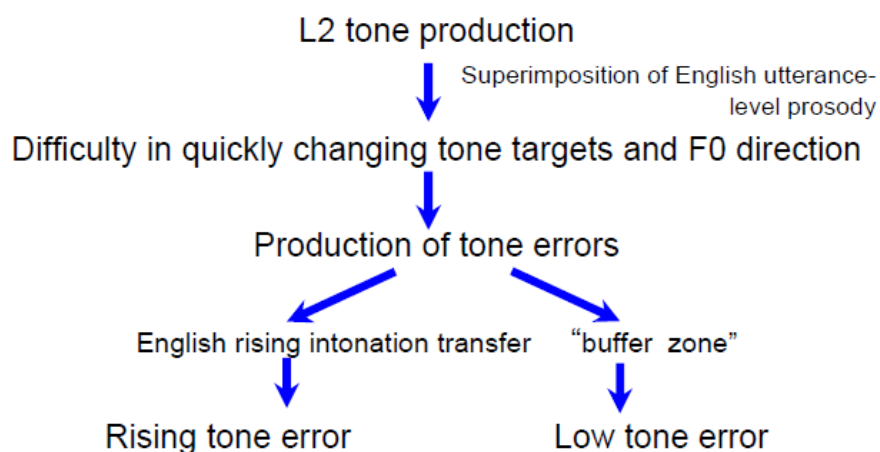


Abbildung 3.2: Das vorläufige Modell des Tonerwerbs nach Yang (2011: 166)

Im Vergleich zu den zahlreichen Studien, die sich mit dem Erwerb von Tönen im SC durch englische Muttersprachler befassen, existieren kaum Studien, welche die Produktion von deutschen Lernern behandeln: z.B. Ding et al (2012) und Ding

(2012). In diesen beiden Studien wird der Zusammenhang von Produktion und Perzeption ein- und zweisilbiger Wörter im SC bei deutschen Lernern unterschiedlicher Niveaustufen getestet. Die Produktion einsilbiger Wörter zeigt bei deutschen Lernern sowohl bei dem Tonregister als auch den Tonkonturen Diskrepanzen. Ding stellt lediglich fest, dass die Produktion zweisilbiger Wörter ein anderes Muster zeigt als die Produktion einsilbiger Wörter. Die Probanden realisieren zweisilbige Wörter jeweils mit einer von ihnen bevorzugten Kontur aus ihrer L1. Jedoch fehlen den zwei Studien sowohl eine systematische Fehleranalyse der Tonproduktion deutscher Lerner als auch eine Untersuchung möglicher Fehlerquellen.

Resümee

Die vorgenannten Studien konzentrieren sich auf die verschiedenen linguistischen Ebenen, auf denen die Töne vorkommen. Sie zeigen, dass Englisch- und Deutsch-Muttersprachler grundsätzliche und systematische Schwierigkeiten bei der Produktion der Töne im SC haben. Unter den vier vollen lexikalischen Tönen stellen T2 und T3 die größte Herausforderung für Lerner dar. Zwischen T2 und T3 bzw. T1 und T4 kommt es häufig zu Verwechslungen. Wenn Töne in einem zweisilbigen Wort produziert werden, treten in der Regel mehr Fehler in der ersten als in der zweiten Silbe auf. Die Produktion des Neutraltons wird in den genannten Studien generell nicht berücksichtigt.

In den vorherigen Studien werden Probanden vom Anfänger- bis zu Fortgeschrittenenniveau untersucht. Dies führt zu unterschiedlichen Schlussfolgerungen über den Erwerb der Töne im SC und weist gleichzeitig darauf hin, dass die Lernersprache ein individuelles, sich jedoch ständig entwickelndes System ist.

Die Fehleranalyse in den zuvor beschriebenen Studien folgt meistens dem Muster des Fehlerzählens. Das heißt, es wird anhand der Beurteilung von Muttersprachlern gezählt, wie viel Prozent der produzierten Töne in ihrer isolierten Form sowie in verschiedenen Position eines Wortes „korrekt“ sind. Dies wirft die

Frage auf, wie zuverlässig die Beurteilung durch Muttersprachler ist. Zudem werden Töne im zusammenhängenden Kontext in den vorherigen Arbeiten stets getrennt betrachtet. Keine der Arbeiten hat den Fokus auf den Übergang zwischen zwei Tönen in Sequenz gelegt.

4 Projektbeschreibung

Im Anschluss an die theoretische Übersicht wird in diesem Kapitel das Projekt der vorliegenden Arbeit beschrieben. Zuerst werden die Forschungsfragen und Hypothesen erläutert. Anschließend werden der neue methodische Ansatz der vorliegenden Arbeit, die theoretische Grundlage und die Probanden beschrieben.

4.1 Forschungsfragen und Hypothesen

Die Forschungsübersicht zu den bisherigen ChaF-Studien zeigt, dass englischsprachige ChaF-Lerner grundsätzlich und systematisch Schwierigkeiten bei der Tonproduktion haben, insbesondere bei Tonsequenzen. Zusätzlich zeigen diese Studien ein relativ einheitliches Bild des Tonerwerbs im SC bei Lernern, deren L1 keine Tonsprache ist. Solche Lerner produzieren Töne in Zitierform mit höherer Genauigkeit als Töne in Sequenz. Fehler treten insbesondere in der ersten Silbe von zweisilbigen Wörtern auf. Zudem kommt es sowohl in der Produktion als auch in der Perzeption oftmals zu Verwechslungen von T2 und T3. Jedoch werden in den genannten Studien nicht die zugrundeliegenden Mechanismen erforscht, die Lerner bei der Tonproduktion einsetzen. Einige Autoren (z.B. Chen 2000 und Yang 2011) argumentieren in ihren Arbeiten, dass die von Lernern produzierten Fehler durch einen negativen Einfluss der L1 (in diesem Fall Englisch) verursacht werden. Jedoch ist unklar, welchen Einfluss die L1 bei der Tonproduktion im SC genau ausübt. Es stellt sich die Frage, ob dieser Einfluss rein linguistischer Natur ist oder der Unterricht ebenfalls eine Rolle spielt.

Zahlreiche Studien liefern bereits Belege dafür, dass der Sprachunterricht bei der Aussprachekompetenz eine nicht zu unterschätzende Rolle spielt (z.B. Elliott 1997; Derwing et al. 1997; Derwing & Rossiter 2003; Dahmen 2013; Wang, Jongman & Sereno 2003; He & Wayland 2010). Diesem Faktor wird in den oben genannten Studien allerdings keine große Bedeutung beigemessen.

Die Untersuchungen der vorliegenden Arbeit sollten zur Beantwortung der folgenden Fragen beitragen:

- (1) Welche Mechanismen verwenden deutsche ChaF-Lerner bei der Realisierung von Tonsequenzen?
- (2) Welche Rolle spielt der bisher praktizierte ChaF-Unterricht?

Die Untersuchungen sollen einerseits zu einem besseren Verständnis der Mechanismen führen, die der Tonproduktion deutschsprachiger ChaF-Lerner bei der Realisierung zweisilbiger Wörter mit zwei Volltönen bzw. einem Voll- und einem Neutralton zugrunde liegen. Andererseits bildet die Anwendung der gewonnenen Erkenntnisse auf die Praxis im Bereich des ChaF-Unterrichts einen weiteren Schwerpunkt der vorliegenden Arbeit. Hier soll gezeigt werden, dass ein explizites Training der Tonsequenzen in verschiedenen großen Domänen weitaus effektiver wäre als ein Training, das sich nur auf die Produktion der lexikalischen Töne in Zitierform beschränkt. Der Fokus der Arbeit liegt daher auf der Gewinnung empirischer Erkenntnisse in der L2-Phonologie und ihrer Anwendung in der Aussprachedidaktik.

In der vorliegenden Arbeit werden zwei separate, empirische Studien durchgeführt. Die erste Studie beschäftigt sich mit der Tonproduktion zweisilbiger Wörter mit zwei vollen Tönen. Die zweite Studie behandelt die Tonproduktion zweisilbiger Wörter mit einem Neutralton. Aus den folgenden Punkten werden die Hypothesen der vorliegenden Arbeit aufgestellt: Aktuelle Erkenntnisse zum Tonsystem im SC und dem prosodischen System im Deutschen, Studien zum Tonerwerb englischsprachiger ChaF-Lerner und Beobachtungen der aktuellen phonetischen Übungspraxis im ChaF-Unterricht.

In der ersten Studie wird tonale Koartikulation zweisilbiger Wörter mit vollen Tönen untersucht und Hypothesen H1 bis H3 werden überprüft:

- H1: Die tonale Realisierung von Tonsequenzen durch deutsche ChaF-Lerner zeigt eine vorwiegend regressive koartikulatorische Tendenz, während chinesische Muttersprachler eine vorwiegend progressive Tendenz zeigen.
- H2: Deutsche ChaF-Lerner versuchen, Töne in Tonsequenzen möglichst klar voneinander abzugrenzen, um die erlernte Zitierform der lexikalischen Töne beizubehalten. Demzufolge zeigen sie bei tonaler Koartikulation eine dissimilatorische Tendenz.
- H3: Explizit gelernte Tonsequenzen, wie beispielsweise die T3-T3 Sequenz, werden von deutschen ChaF-Lernern und chinesischen Muttersprachlern ähnlich realisiert.

Der Neutralton wird aufgrund seines viel debattierten Status im SC stets im Forschungsbereich Spracherwerb vernachlässigt (Yang 2010). Die zweite Studie versucht daher erstmals, anhand der folgenden Hypothese einen Einblick in die Produktion des Neutraltons von ChaF-Lernern zu gewinnen:

- H4: Lerner produzieren den Neutralton im SC wie eine unbetonte Silbe im Deutschen, wodurch lexikalische Kontraste in der Tonproduktion zum Teil verloren gehen und die tonale Realisierung bestimmter Tonpaare miteinander verschmilzt.

Die oben genannten Hypothesen werden in den Kapiteln 5 und 6 detailliert erläutert und zur Diskussion in Kapitel 7 noch einmal aufgegriffen.

4.2 Produktion und Perzeption der Töne

Der Fokus der vorliegenden Arbeit liegt auf der Produktion der Töne im SC durch deutsche ChaF-Lerner. Ein Perzeptionstest wird im Rahmen der Untersuchung nicht durchgeführt, weil das Verhältnis zwischen Produktion und Perzeption der

Töne nicht eindeutig ist. Verschiedene Studien haben bis jetzt zu unterschiedlichen Konklusionen über das Verhältnis zwischen Produktion und Perzeption der Töne geführt. Leather (1990) stellt einen signifikant engen Zusammenhang zwischen Fehlern bei der Produktion und Perzeption bei Nicht-Muttersprachlern fest. Elliot (1991) entdeckt jedoch keine enge Relation zwischen der Produktion und der Perzeption, insbesondere bei T3 und T4. Chen (1997) konnte ebenfalls keine Korrelation zwischen den Fehlern in der Produktion und Perzeption feststellen. Die Erklärung dafür könnte sein, dass die Produktion durch einen kontinuierlichen, akustischen Parameter – F0 dargestellt wird, während die Perzeption der Töne verschiedene Wahrnehmungskategorien aufweist. Abbildung 4.1 illustriert den Unterschied zwischen Produktion und Perzeption von Tönen bei Muttersprachlern des SC (Yang 2010).

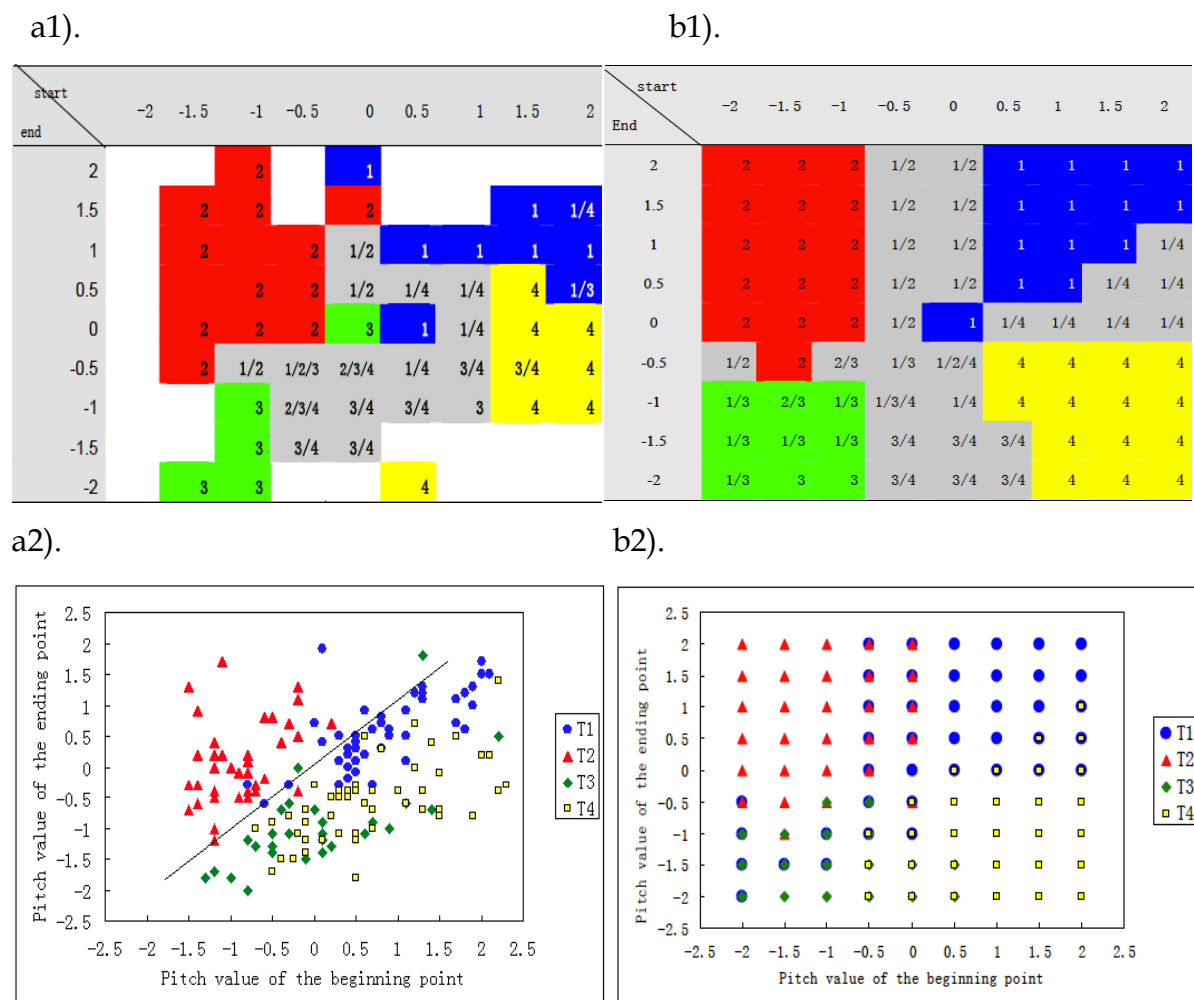


Abbildung 4.1: Vergleich zwischen Produktion (linke Spalte, a1 und a2) und Perzeption (rechte Spalte b1 und b2) bei Muttersprachlern des SC (Yang 2010: 98)

In a1) und a2) der Abbildung 4.1 sieht man die Verteilung der vier lexikalischen Töne innerhalb des muttersprachlichen Produktionsspektrums. In beiden Bildern werden die Onset- und Offset-Werte von Tönen graphisch dargestellt. Bis auf T2 (in a2 rote Dreiecke) existieren zwischen den anderen Tönen keine eindeutigen Grenzen. In b1) und b2) werden die Kategorien innerhalb des muttersprachlichen Perzeptionsspektrums dargestellt. Im Gegensatz zum Produktionsspektrum sind die perzeptiven Kategorien der Töne relativ klar voneinander abgegrenzt (siehe Farbblocke in b2)).

In den meisten ChaF-Studien werden Fehler in der Produktion anhand der kategorischen Perzeption der Töne durch Muttersprachler festgestellt. Die Fehler in der Perzeption werden hingegen anhand der tonalen Kategorien der Lerner erfasst. Um eine eindeutige Korrelation zwischen der Produktion und der Perzeption der Töne bei Lernern aufzuzeigen, müsste vorausgesetzt sein, dass Muttersprachler und Lerner die gleichen tonalen Kategorien besitzen. Dies wird jedoch eindeutig in der Studie von Yang (2010) widerlegt. Hinzu kommt, dass die Perzeption dem Effekt des so genannten „muttersprachlichen Filters“ unterliegt. Trubetzkoy (1939) wurde auf dieses Phänomen aufmerksam, welches das phonologische System der L1 eines Lerners als ein „Sieb“ beschreibt. Bei der Perzeption wird alles, was der L1 nicht zuzuordnen ist, durch dieses „Sieb“ gefiltert, wodurch die Wahrnehmung bestimmter Strukturen, die in der L1 nicht vorhanden sind, erschwert oder sogar verhindert wird. Lexikalische Töne sind in der deutschen Sprache nicht vorhanden, weswegen anzunehmen ist, dass deutsche Lerner Fehler bei der Perzeption der Töne machen. Da das Ziel der vorliegenden Arbeit nicht darin liegt, die Relation zwischen Produktion und Perzeption der Töne zu erforschen, enthält sie keine perzeptive Untersuchung. Des Weiteren benötigen Perzeptionsstudien eine andere wissenschaftliche Methodik und sollten in einem separaten Rahmen angestellt werden.

4.3 Methodik

Der methodische Ansatz der vorliegenden Arbeit unterscheidet sich stark von den in Kapitel 3 vorgestellten Studien im Bereich ChaF. Wie in Abschnitt 4.1.3. erwähnt, basiert die Fehlerermittlung in den meisten ChaF-Studien auf der perzeptiven Beurteilung durch Muttersprachler. Es werden hauptsächlich zwei Arten von Tests durchgeführt: Diskriminierungs- und Identifikationstests. In einem Diskriminierungstest sollen sich die Muttersprachler bei einer Tonproduktion entscheiden, ob es sich um eine „korrekte“ Produktion handelt. In einem Identifikationstest werden die Beurteiler darum gebeten, die gehörten Töne in der L2-Produktion zu identifizieren. Beide Verfahren können jedoch das grundsätzliche Problem nicht umgehen, dass eine Tonrealisierung von Lernern meistens nicht eindeutig genug für Muttersprachler zu beurteilen ist, weil jeder Muttersprachler individuelle Tonkategorien besitzt. Eine muttersprachliche Beurteilung der Tonproduktionen in der L2 bedeutet, dass ein von Lernern produziertes akustisches Kontinuum in die Tonkategorien der Muttersprachler eingeordnet werden muss. Dies führt zu sehr variablen Ergebnissen innerhalb einer Gruppe von Beurteilern. In den vergangen Studien versuchen die Autoren eine Lösung für das Problem durch eine Erhöhung der Anzahl von Beurteilern zu finden, z.B. Wang et al. (2003), oder ein „*inter-rater judgement*“ Verfahren zu verwenden, z.B. Shen (1989), Sun (1998), Chen (2000) oder Hao (2012). In diesem Verfahren beurteilen mehrere Muttersprachler gleichzeitig die Tonproduktion von Lernern, und bei Unstimmigkeiten wird die Entscheidung nach dem Mehrheitsprinzip getroffen. Trotz solcher Maßnahmen bleibt es nahezu unmöglich, die Tonproduktion von Lernern als absolut richtig oder falsch einzustufen oder definitiv einer der vier Tonkategorien zuzuordnen. Es lässt sich bestenfalls feststellen, ob eine Tonproduktion der muttersprachlichen Produktion näher oder weiter entfernt ist. Eine Analyse, die ausschließlich auf muttersprachlichen Beurteilungen basiert, ist deswegen nicht hinreichend aussagekräftig. Aus diesem Grund wird der Fokus in der vorliegenden Arbeit auf die tatsächliche Realisierung gelegt. Die Analyse der beiden Studien basiert auf gemessenen Parametern der Tonproduktion, wie der F0 oder der Dauer.

4.4 Zugrundeliegende Theorien und Modelle des Spracherwerbs

Die Untersuchungen der vorliegenden Arbeit basieren hauptsächlich auf der „*Interlanguage Hypothesis*“ (Selinker 1972; Major 2001, siehe auch Kapitel 3). Diese nimmt an, dass sich die Lernersprache (*Interlanguage*) aus Elementen der L1, L2 sowie den aus dem Erwerbsprozess entstandenen Eigenschaften zusammensetzt. In den vorliegenden Untersuchungen werden daher L1, L2 und die Lernersprache der Probanden berücksichtigt. Die aufgenommene L2-Produktion der Lerner wird unter Berücksichtigung der Fachliteratur über das Deutsche analysiert, um diejenigen Eigenschaften der Lernersprache zu ermitteln, die auf L1 zurückzuführen sind. In die Analyse wird zusätzlich die Auswirkung der Unterrichtspraxis im Bereich ChaF einbezogen.

Auch die Theorie des negativen Transfers ist eine wichtige Theoriegrundlage der vorliegenden Arbeit. In den Untersuchungen werden linguistisch kontrastive Analysen des SC und der deutschen Sprache durchgeführt, um die Hypothesen aufzustellen.

4.5 Probanden

Im Folgenden werden zuerst die Teilnehmer der muttersprachlichen Kontrollgruppe und anschließend die deutschen Probanden beschrieben. Alle Informationen zu den deutschen Probanden wurden mittels eines Fragebogens erfasst (ein beispielhafter Fragebogen befindet sich in Appendix A).

4.5.1 Muttersprachliche Kontrollgruppe

Wie in Kapitel 2 beschrieben, basiert die Aussprachenorm des SC auf dem Dialekt von Beijing. Die für die vorliegende Arbeit eingesetzten Muttersprachler sprechen ausschließlich diese Standardaussprache des SC. Sie sind alle in Beijing geboren und aufgewachsen. Insgesamt nehmen 16 Muttersprachler teil (zehn an der ersten

und sechs an der zweiten Studie). Sie sind zwischen 25 und 45 Jahre alt, leben zur Zeit der Aufnahmen in Deutschland und haben Hochdeutsch als Fremdsprache gelernt. Die Sprecher aus der muttersprachlichen Kontrollgruppe werden in dieser Arbeit mit „CM“ abgekürzt.

4.5.2 Deutsche Probanden

Insgesamt haben sechzehn deutschen Probanden an den Studien teilgenommen (zehn an der ersten und sechs an der zweiten Studie). Die meisten von ihnen kommen aus dem Großraum Köln, sind zwischen 21 und 27 Jahre alt und studieren an der Universität zu Köln. Die deutschen Lerner werden in dieser Arbeit als „DL“ bezeichnet.

4.5.2.1 Lernsituation der deutschen Probanden

Die deutschen Probanden haben SC in erster Linie als Fremdsprache gelernt. Bis auf zwei von ihnen haben alle Lerner im Rahmen eines Studiums an der Universität zu Köln fünf Semester mit durchschnittlich acht bis zehn Semesterstunden Unterricht pro Woche SC gelernt. Die beiden anderen Probanden haben SC jeweils 6 bzw. 7 Semester studiert. Die meisten Probanden geben an, dass sie neben dem Sprachunterricht im universitären Umfeld wöchentlich 30 bis 60 Minuten Chinesisch mit Muttersprachlern sprechen. Zur Zeit der Aufnahmen haben alle Probanden bereits einen Aufenthalt in China absolviert, dessen Dauer von einem bis zu 12 Monaten variiert. Während des Aufenthaltes erhielten die meisten Probanden in der Regel weiterhin ChaF-Unterricht an einer Hochschule ihres Aufenthaltsortes und hatten regelmäßigen Kontakt zu Muttersprachlern. Der Spracherwerb der deutschen Probanden verlief dementsprechend größtenteils institutionell gesteuert, mit unterschiedlich langen Phasen eines zusätzlichen natürlichen, ungesteuerten Spracherwerbs. Wichtig für die vorliegende Arbeit ist, dass beide Komponenten im Prozess des Erlernens jedes Probanden enthalten sind.

4.5.2.1 Lernerspezifische Faktoren der Probanden

In Form eines Fragebogens werden die lernerspezifischen Faktoren der deutschen Probanden erfasst, wie das Alter beim ersten Sprachkontakt, die Länge des Aufenthaltes in China und der Umfang des Sprachgebrauchs außerhalb der Universität (siehe Tabelle 4.1). Das Alter beim ersten Sprachkontakt liegt bei allen deutschen Proband im Erwachsenenalter (außer DL7 und DL9, die schon mit 16 Jahren Chinesisch-Unterricht nahmen). Somit haben sie ähnliche Voraussetzungen für die erreichbare Kompetenz der Aussprache.

Nr.	m /f	Alter	Alter bei Beginn des Erwerbs	Anzahl der Semester	Sprachge- brauch pro Woche außer Universität (Min.)	Selbst- einschät- zung ¹⁶	Aufenthalt in China (Monaten)
DL1	f	25	21	5	180	1>3>2>4	12
DL2	f	22	19	7	60-70	3>2>4>1	12
DL3	m	22	19	5	60	4>1>2>3	12
DL4	m	23	20	5	90	2>3>1>4	1
DL5	f	22	18	5	0	2>1>3>4	12
DL6	f	23	19	5	10	2>4>3>1	12
DL7	m	24	16	6	10	3>2>4>1	10
DL8	f	22	20	5	30	4>2>3>1	6
DL9	f	23	16	5	60	2>3>4>1	9
DL10	f	21	19	5	30-60	3>4>2>1	1,5
DL11	f	25	20	5	400 ¹⁷	3>2>1>4	12
DL12	f	23	20	5	20	2>3>4>1	6
DL13	m	24	20	5	0-10	2>4>3>1	12
DL14	f	23	19	5	30	3>2>4>1	10
DL15	m	22	19	6	60	2>3>4>1	12
DL16	m	27	23	7	0-10	2>1>3>4	12

Tabelle 4.1: Übersicht der Lerner-Informationen. DL1 – DL10 nehmen an der ersten und DL11 – DL16 an der zweiten Studie teil.

¹⁶ In der Rubrik „Selbsteinschätzung“ sollten die Probanden die vier lexikalischen Töne nach ihrem individuell empfundenen Schwierigkeitsgrad einordnen. „1“>“2“ bedeutet, dass T1 schwieriger als T2 eingestuft wird.

¹⁷ Diese Probandin arbeitet zur Zeit der Aufnahme in einem China Restaurant und gibt ihre Arbeitszeit als die Zeit des Sprachgebrauchs an.

Die Probandengruppe ist auch insofern homogen, als dass alle Probanden von den gleichen Dozenten und mit dem gleichen Lehrmaterial unterrichtet werden. Der Fragebogen beinhaltet außerdem Angaben zur Muttersprache und weiteren erlernten Fremdsprachen. Alle Probanden haben im Laufe des Lebens zusätzlich Englisch als Fremdsprache gelernt. Vierzehn der sechzehn Lerner haben zusätzlich Französisch gelernt. Keiner der Probanden hat neben SC jemals eine weitere Tonsprache gelernt. Ursprünglich haben achtzehn Lerner den Fragebogen ausgefüllt, doch zwei Kandidaten konnten nicht berücksichtigt werden, weil sie bilingual aufgewachsen sind und neben Deutsch auch Kantonesisch sprechen. Insgesamt stellen sich die Probanden bezüglich der lernerspezifischen Faktoren als relativ homogen dar.

4.5.3 Die Evaluierung der fremdsprachlichen Kompetenz

Seit 2001 liegt allen Fremdsprachenlehrern in Europa der *„Gemeinsame Europäische Referenzrahmen für Fremdsprachen“* vor, in dem die Komplexität des Fremdspracherwerbs in Niveaustufen und handlungsorientierten Kompetenzen strukturiert ist (Guder 2005). Das wesentliche Ziel des Referenzrahmens liegt unter anderem darin, eine erfolgreiche Beschreibung der Kompetenzniveaus zu unterstützen. Dieser Referenzrahmen beschränkt sich allerdings ausschließlich auf Sprachen des westlichen Zweigs der indoeuropäischen Sprachen (Guder 2005: 63). SC gehört nicht zu dieser Sprachfamilie und wird deshalb nicht in dem Referenzrahmen berücksichtigt (Guder 2005: 64). Seit einigen Jahren existiert in China eine entsprechende Leitlinie der staatlichen standardisierten Prüfung zum Nachweis von Chinesisch-Kenntnissen - „*HSK*“ (汉语水平考试 *Hanyu Shuiping Kaoshi*). Dieses Evaluierungssystem betrachtet allerdings die inländischen Chinesisch-Intensivkurse als Standardlernform, d.h. Lerner nehmen in China an ChaF-Unterricht teil. Nach dem neusten Stand der HSK-Leitlinie ist speziell für die mündlichen Prüfungen die folgende Anzahl an Unterrichtseinheiten zur Erreichung der entsprechenden Kompetenzstufen vorgeschrieben, wie Tabelle 4.3 darstellt:

Neue HSK Leitlinie Mündliche Prüfung	Unterrichtseinheiten (UE)	Wortschatz
Grundstufe	2-3 UE pro Woche 1-2 Semester	ca. 200
Mittelstufe	2-3 UE pro Woche 2-4 Semester	ca. 900
Fortgeschrittene Stufe	2-3 UE pro Woche über 4 Semester hinaus	ca. 3000

Tabelle 4.2: Neue HSK-Leitlinie der mündlichen Prüfung (adaptiert vom „*Bulletin of HSK Test*“ 2011).

Für die vorliegende Arbeit sind die Europäischen Referenzrahmen als Einstufungsinstrument nicht einsetzbar, weil diese ausschließlich europäische Sprachen berücksichtigen. Eine Einstufung anhand der HSK-Leitlinie wird lediglich bezüglich der Wortschatzgröße und der Anzahl der UE vorgenommen. Eine genaue Umrechnung in UE an deutschen Hochschulen ist allerdings bis dato nicht bekannt. Mangels anderer Messinstrumente werden die Probanden auf Basis der Wortschatzgröße als fortgeschritten eingestuft, obwohl derartige Einstufungen bezüglich der Aussprachekompetenz relativ wenig aussagkräftig sind.

5 Untersuchung der tonalen Koartikulation

Im Folgenden wird die Realisierung von Tonsequenzen in zweisilbigen Wörtern untersucht. Das Experiment behandelt zweisilbige Wörter produziert in Zitierform und innerhalb von Äußerungen.

5.1 Hypothesen

Wie schon in Kapitel 2 erwähnt, unterliegen Höhe und Konturen der lexikalischen Töne den koartikulatorischen Effekten, wenn diese nacheinander produziert werden. Tonale Koartikulation im SC sowie in ostasiatischen Sprachen ist ausführlich erforscht (z.B. Han & Kim 1974; Brunelle 2009; Abramson 1979; Gandour et al. 1994; Gandour et al. 1996; Peng 1997; Shih 1988; Shen 1990; Xu 1994, 1997). Jedoch widmet sich keine der Studien einer systematischen Untersuchung zur tonalen Koartikulation in der Produktion von Lernern, deren L1 keine Tonsprache ist. Die Mechanismen, auf die Lerner bei der tonalen Produktion im SC zurückgreifen, sind kaum bekannt. Sich mit dem Erwerb der Töne im SC befassende Studien zeigen, dass isolierte Töne mit höherer Genauigkeiten produziert werden als hintereinander auftretenden Töne (z.B. He 2010). Zudem postulieren einige Autoren, dass unterschiedliche Wortakzentmuster sowie die Intonation der L1 die Tonproduktion von Lernern negativ beeinflussen (White 1981, Chen 2000, Yang 2011, unter anderen).

Es stellt sich die Frage, welche zugrundeliegenden Mechanismen hinter dem bereits berichteten Phänomen stecken, dass Lerner erhebliche Schwierigkeiten haben, sobald sie die Töne in Sequenzen statt in Zitierform produzieren müssen und welche Rolle der ChaF-unterricht dabei spielt. Die vorliegende Studie versucht diese Fragen zu beantworten.

Der Konsens der unterschiedlichen Studien zu tonaler Koartikulation im SC ist, dass der überwiegende koartikulatorische Effekt progressiv ist. Es existiert jedoch auch im bestimmten tonalen Kontext geringfügig regressive Koartikulation. Wenn

ein beliebiger Ton vor einem Ton mit einem niedrigen Onset (T2 oder T3) steht, ist die Kontur des betreffenden Tons höher, als wenn dieser vor einem Ton mit einem hohen Onset steht.

Abbildung 5.1 zeigt ein Beispiel für die progressive tonale Koartikulation im SC. T1 wird in der Silbe /ma/ in der Anfangsposition (im folgenden: P1) oder in der Endposition (im folgenden: P2) im zweisilbigen Wort /mama/ produziert.

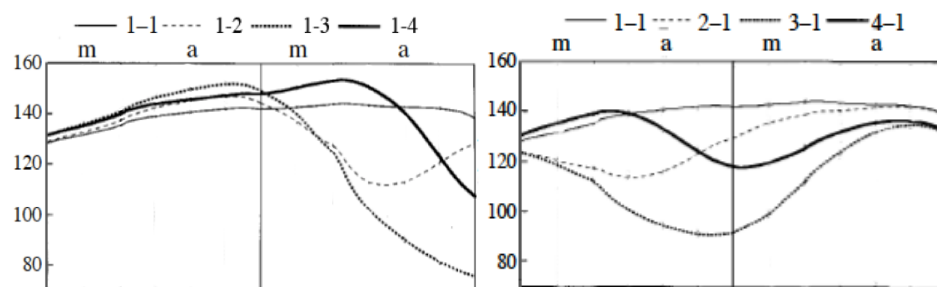


Abbildung 5.1: Beispiel der progressiven Koartikulation anhand der Realisierung von T1 in P1 und P2 eines zweisilbigen Wortes (Xu 1997: 69 und 75).

Wenn T1 in P1 produziert wird, zeigt er ungeachtet der Identität der nachfolgenden Töne sehr wenig Variation am Offset. Wird derselbe Ton in P2 produziert, wird der Onset-Wert des Tons vom Offset-Wert der vorausgehenden Töne stark beeinflusst und zeigt demzufolge viel Variation in diesem Bereich. Nach Flemming (2011) unterliegen die sprachspezifischen Muster in der tonalen Koartikulation hauptsächlich dem Einfluss des konkreten Tontargets in der betroffenen Sprache. Um die tonalen koartikulatorischen Muster in der Produktion von Muttersprachlern und Lernern zu untersuchen, müssen als erstes die Tontargets und ihre Alignierungen im SC und Deutschen beschrieben werden.

5.1.1 Tonale Alignierung im Standardchinesischen

Die oben dargestellte relativ variable F0-Kontur eines lexikalischen Tons, der in Sequenz produziert wird, führt zu der Frage, was genau die Targets der Töne im SC sind und wie sie phonetisch realisiert werden. Xu & Wang (2001) schlagen am Beispiel des SC einen allgemeinen Rahmen vor, der die oberflächliche F0-

Variation in der tonalen Implementierung generell erklären soll. Demnach können Tontargets in statische und dynamische Targets unterteilt werden. Statische Targets haben eine Spezifikation des Registers, z.B. [H] (*high* - hoch) oder [L] (*low* - niedrig), während dynamische Targets eine Spezifikation der linearen Bewegung haben, z.B. [R] (*rising* - steigende Bewegung) oder [F] (*falling* - fallende Bewegung). Folgende Implementierungsregeln der Tontargets wurden von den beiden Autoren aufgestellt:

[1] *A pitch target is implemented in synchrony with the host , i.e. , starting at its onset and ending at its offset .*

[Die Implementierung eines Tontargets ist (zeitlich) synchronisiert mit der tontragenden Domäne, d.h., sie startet an ihrem Onset und endet an ihrem Offset.]

(Xu & Wang 2001: 322, meine Übersetzung)

Die Regel [1] sagt aus, dass die Implementierung eines Tontargets erst gestartet wird, wenn das vorausgehende Target und die Grenze der tontragenden Domäne schon erreicht sind.

[2] *Throughout the duration of the host , the approximation of the pitch target is continuous and asymptotic .*

[Während der gesamten Dauer der tontragenden Domäne ist die Annäherung des Tontargets kontinuierlich und asymptotisch.]

(Xu & Wang 2001: 322, meine Übersetzung)

Asymptotische Annäherung bedeutet, dass eine F0-Bewegung bei der Annäherung an ein Tontarget zuerst schnell und im Verlauf der Zeit verlangsamt erfolgt. Am Ende der tontragenden Domäne ist die maximale Annäherung an das

Tontarget erreicht. Demnach ist die Silbengrenze im SC ein Referenzpunkt, an dem sich die tonale Implementierung orientiert. Das Onset einer Silbe müsste Raum für den Übergang vom vorausgehenden Tontarget zum Beginn der Annäherung an das nächste lassen. In Abbildung 5.2 erkennt man eine große Abweichung am Anfang der Silben bei beiden Targets aufgrund des Einflusses des vorausgehenden Targets. Die maximalen Annäherungen findet man an den Silbengrenzen.

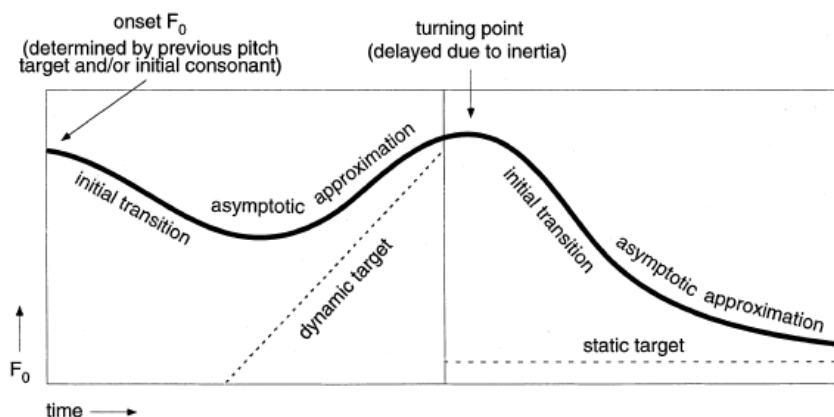


Abbildung 5.2.: Schematische Darstellung von hypothetischen Tontargets (gestrichelte Linien) und der oberflächlichen Kontur der Realisierung (durchgezogene Linien). Die vertikalen Linien repräsentieren Silbengrenzen zwei sequentiell tontragender Silben (Xu & Wang 2001: 321).

Nach Flemming (2011) kann der sprachspezifische Zeitpunkt, wann der Übergang zwischen zwei Tontargets stattfindet, von zwei Faktoren beeinflusst werden: der Typ des Tontargets (z.B. statisch oder dynamisch) und die Position der relevanten Teile der Tontargets (z.B. Silbenanlaut oder der Reim). Zhang (2004) findet Beweise dafür, dass der Reim einer Silbe wichtig für den tonalen Kontrast ist. Da der Reim in Silben generell die höchste Intensität und Periodizität hat, wird ein Ton an dieser Stelle am besten wahrgenommen. Flemming (2011) postuliert, dass die überwiegend progressive tonale Koartikulation in den verschiedenen Tonsprachen aufgrund dieser Eigenschaft entsteht. Der Reim in Silben ist für den tonalen Kontrast wichtiger als das Onset, weshalb der Übergang zwischen zwei Tontargets in vielen Sprachen oft im Bereich des Onsets stattfindet. Brunelle (2009) stellt beispielsweise fest, dass das Onset der vietnamesischen Töne eine kleinere Rolle bei der Tonidentifikation spielt als das Offset.

5.1.2 Tonale Alignierung im Deutschen

Tonale Spezifikationen (z.B. Tonakzente) sind in Sprachen wie Deutsch mit bestimmten Segmenten als eine phonologisch abstrakte Eigenschaft verankert. Tonale Alignierung dagegen ist eine phonetische Eigenschaft und stellt die zeitliche Koordination zwischen den Ereignissen in der F₀-Kontur und bestimmten Segmenten dar. Diverse Studien beschäftigen sich mit den akustischen Details der Intonationskonturen und zeigen eine stabile und vorhersagbare Alignierung zwischen Wendenpunkten in der F₀-Kontur und bestimmten Segmenten (Ashby 1978; Silvermann & Pierrehumbert 1990; Prieto et al. 1995; Arvaniti & Ladd 1995; Arvaniti et al. 1998;). Zum Beispiel wird der Gipfel des Pränuklearakzents früher in der Silbe aligniert, wenn sich die Silbenanzahl zwischen diesem Pränuklearakzent und dem nachfolgenden Nuklearakzent oder Grenzton verkleinert. Das heißt, der rechts stehende Nuklearakzent oder Grenzton übt einen rückwärtsgerichteten Zeitdruck auf die Alignierung des Gipfels eines Pränuklearakzents aus (Silvermann & Pierrehumbert 1990; Prieto et al. 1995). Wenn aber die Silbenanzahl zwischen einem Pränuklearakzent und dem nächsten Akzent oder Grenzton konstant bleibt und nur die tatsächliche Dauer sich ändert, wird festgestellt, dass der Gipfel konstant in der Nähe des Anfangs der unbetonten Silbe oder unmittelbar danach aligniert wird (Ladd 2008). Zum Beispiel zeigen Arvaniti, Ladd und Mennen (1998), dass im Griechischen der Anfang einer steigenden Kontur in einem Pränuklearakzent immer am Onset oder ganz kurz vor der akzentuierten Silbe aligniert ist. Dabei spielt die Dauer und Auslenkung der Kontur keine Rolle, wie in der Abbildung 5.3 illustriert:

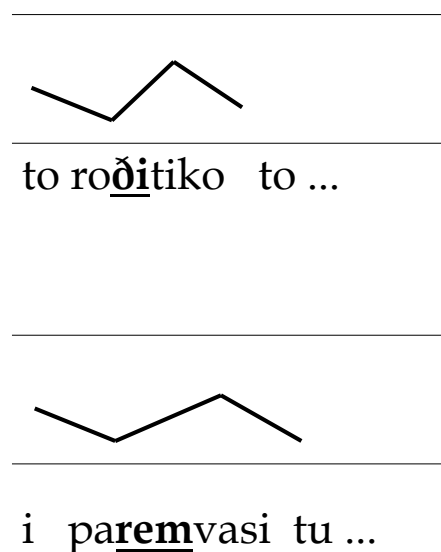


Abbildung 5.3: Stilisierte Tonhöheverläufe, um „segmentale Verankerungen“ darzustellen. Die Wendepunkte am Anfang und Ende des akzentuierten Anstiegs werden mit dem Anfang der akzentuierten Silbe „-đi“ bzw. „-rem“, und dem Anfang der nachfolgenden nicht-akzentuierten Silbe „-ti“ bzw. „-va“ aligniert, Ladd (2008: 175)

Die Ergebnisse der oben genannten Studien zeigen, dass die F0-Kontur als solche keine fundamentale Einheit bei der tonalen Realisierung in Sprachen wie Deutsch bildet, sondern nur als Transition zwischen den Targets fungiert. Bei der tonalen Realisierung im Deutschen scheint der Silbenanfang ein wichtiger Referenzpunkt zu sein, während für lexikalische Töne im SC das Silbenende der wichtigste Referenzpunkt ist. Das Ergebnis einer neuro-linguistischen Studie von Kaan und Kollegen (2008) zeigt zusätzlich, dass ein Sprecher einer Nicht-Tonsprache (in diesem Fall Englisch) bei der Tonwahrnehmung sensibler auf die Änderung der F0-Werte am Onset reagiert, als ein Sprecher einer Tonsprache (in diesem Fall Thailändisch). Der F0-Wert am Onset eines Tons scheint bei der Tonidentifikation für einen Nicht-Tonsprachensprecher eine größere Rolle als für einen Tonsprachensprecher zu spielen. Nach Flemming (2011) sind unterschiedliche Muster oder Ausmaße der tonalen Koartikulation in verschiedenen Sprachen die Reaktionen auf die sprachspezifischen *constraints*. Daraus ergibt sich die Frage, ob diese Annahme auch auf die tonale Produktion in L2 übertragbar ist. Die unterschiedliche Bewertung des F0-Wertes am Onset und Offset eines Tons von

Muttersprachlern und Lernern führt möglicherweise zu unterschiedlichen Mustern der tonalen Koartikulation. Dies bedeutet, dass Muttersprachler bei der Realisierung einer Tonsequenz, z.B. T3-T1, erst nach Ende der ersten Silbe anfangen, sich an das zweite Tontarget anzunähern wie in Abbildung 5.4 (b) und Lerner schon viel früher anfangen, sich an das Target anzunähern wie in 5.4 (a).

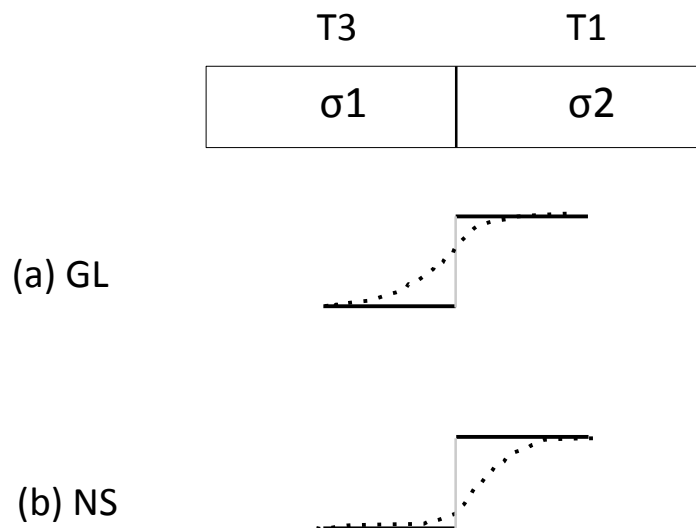


Abbildung 5.4: Hypothetische Schemata der F0-Kontur von Tonsequenz T3-T1, realisiert von deutschen Lernern (DL) und Muttersprachlern (CM).

Demnach scheint die zweite Silbe bei Lernern mehr Einfluss auf die vorausgehende Silbe zu haben, was eher dem regressiven Effekt entspricht. Bei Muttersprachlern hat die erste Silbe mehr Einfluss auf die nachfolgende, was dem progressiven Effekt entspricht, weshalb die folgende Hypothese H1 aufgestellt wird:

- H1: Die tonale Realisierung von Tonsequenzen durch deutsche ChaF-Lerner zeigt eine vorwiegend regressive koartikulatorische Tendenz, während chinesische Muttersprachler eine vorwiegend progressive Tendenz zeigen.

In der vorliegenden Arbeit wird auch die Rolle des Unterrichts untersucht. Wie schon in Kapitel 3 berichtet, werden Übungen zur phonetischen Realisierung von Tonsequenzen im ChaF-Unterricht, den die Probanden dieser Studie besuchten, in der Regel vernachlässigt, weshalb zu erwarten ist, dass ihre Strategien bei der Realisierung von Tonsequenzen durch die Konzentration auf die Produktion isolierter lexikalischer Töne beeinflusst wird. Hingegen wird erwartet, dass eine explizit geübte Produktion von Tonsequenzen von Lernern mit einer zielsprachengerechten Kontur realisiert wird. Daher werden die Hypothesen H2 und H3 aufgestellt:

H2: Deutsche ChaF-Lerner versuchen, Töne in Tonsequenzen möglichst klar voneinander abzugrenzen, um die erlernte Zitiertform der lexikalischen Töne beizubehalten. Demzufolge zeigen sie bei tonaler Koartikulation eine dissimilatorische Tendenz.

H3: Explizit gelernte Tonsequenzen, wie beispielsweise die T3-T3 Sequenz, werden von deutschen ChaF-Lernern und chinesischen Muttersprachlern ähnlich realisiert.

5.2 Methodik

Die Richtung der tonalen Koartikulation hängt eng mit dem Zeitpunkt des Übergangs zwischen zwei benachbarten Tönen zusammen, der zum größten Teil von den sprachspezifischen Tontargets bestimmt wird (Flemming 2011). Ein Übergang von einem hohen zu einem tiefen Tontarget kann, wie in Abbildung 5.5 dargestellt, zu verschiedenen Zeitpunkten starten:

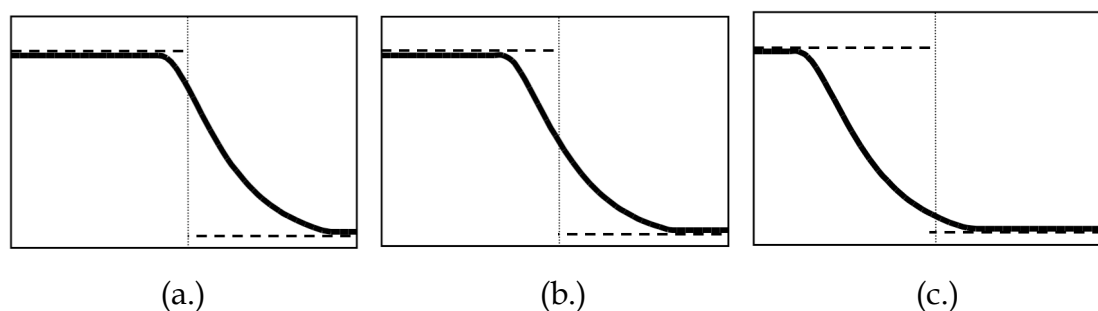


Abbildung 5.5: Stilisierte Übergang von einem hohen zu einem tiefen Tontarget. Die vertikale gepunktete Linie steht für eine Silbengrenze und die horizontale gestrichelte Linie für die Grundform der betreffenden Töne (Flemming 2011)

In (a.) startet der Übergang direkt am Anfang der zweiten Silbe, sodass der F0-Wert am Onset des zweiten Tons beträchtlich vom F0-Wert am Offset des ersten Tons beeinflusst wird und entsprechend eine erhebliche Abweichung zur kanonischen Form aufweist, was dem progressiven Effekt entspricht. In (c.) startet der Übergang dagegen weit vor der Silbengrenze und ist schon am Anfang der zweiten Silbe beendet. Der F0-Wert am Offset des ersten Tons ist stark vom F0-Wert am Onset des zweiten Tons beeinflusst und weicht in diesem Bereich beträchtlich von der kanonischen Form ab, was den regressiven Effekt demonstriert. Als drittes mögliches Muster erstreckt sich der Übergang über die Silbengrenze, wie in (b.) abgebildet. Dieses Muster kann als wechselseitige Koartikulation beschrieben werden. Die Richtung der Koartikulation kann also im Endeffekt durch den Grad der Abweichung von der kanonischen Form zur tonalen Realisierung in relevanten Bereichen (Onset vs. Offset) eines Tons dargestellt werden und ist deshalb ein zuverlässiger Indikator zur Identifizierung der Richtung der tonalen Koartikulation. Um den Grad der Abweichung zu bestimmen, wird in der Produktion von Muttersprachlern und Lernern die **Variabilität** in den betreffenden Bereichen berechnet und verglichen. Die detaillierte Berechnung der Variabilität wird im Abschnitt 5.2.4 beschrieben.

5.2.1 Testmaterial

Es werden sechs einzelne Silben mit jeweils vier lexikalischen Tönen verwendet: *wei*, *mao*, *wu*, *mo*, *wa*, und *mi*. Daraus ergeben sich 24 sinnvolle einsilbige Wörter.

Diese 24 einsilbigen Zielwörter werden isoliert produziert. (5.1) zeigt einige Beispiele:

(5.1)	<i>wēi</i>	[uei ⁵⁵]	微	“winzig”
	<i>wéi</i>	[uei ³⁵]	围	“umzingeln”
	<i>wěi</i>	[uei ²¹⁴]	伟	“großartig”
	<i>wèi</i>	[uei ⁵¹]	位	“Position”

Aus obigen sechs einsilbigen Wörtern werden zweisilbige Zielwörter zusammengesetzt und ergeben 48 sinnlose Komposita. Sie haben alle die segmentalen Strukturen *weimao*, *wumo* und *wami*. Insgesamt tragen die Komposita 16 mögliche tonale Kombinationen der vier lexikalischen Töne. Einige von ihnen werden in (5.2) gezeigt:

(5.2)	<i>wēi māo</i>	[uei ⁵⁵ mau ⁵⁵]	微猫
	<i>wēi máo</i>	[uei ⁵⁵ mau ³⁵]	微毛
	<i>wēi mǎo</i>	[uei ⁵⁵ mau ²¹⁴]	微卯
	<i>wēi mào</i>	[uei ⁵⁵ mau ⁵¹]	微帽

(Vollständige Wortliste siehe Appendix B.)

Als Anlautkonsonanten der Silben werden nur Sonoranten gewählt, um unnötige Turbulenzen im Signal und unerwünschte Blockierungseffekte zu vermeiden, wie Gandour (1994) sie feststellte.

5.2.2 Ablauf der Datenerhebung

Alle Probanden sitzen bei der Aufnahme in der Sprachkabine der Abteilung Phonetik des Linguistischen Instituts der Universität zu Köln. Sie werden aufgefordert, alle Zielwörter laut vorzulesen. Alle Zielwörter werden ihnen in chinesischen Zeichen und *Hanyu Pinyin* gezeigt. Den Probanden wird vor der

Aufnahme Zeit gegeben, sich mit den Zielwörtern vertraut zu machen. Jedes Wort wird dreimal wiederholt. Wenn sie einen Fehler gemacht haben oder unzufrieden mit der eigenen Aussprache sind, können sie jederzeit das betreffende Wort wiederholen. Zusammen werden 8640 Datenpunkte erhoben: 24 einsilbige Einheiten + 48 zweisilbige Einheiten × 3 Wiederholungen × 20 Probanden (10 Muttersprachler und 10 Lerner) × 2 koartikulatorische Richtungen. Von diesen Datenpunkten werden 556 wegen heftiger Laryngalisierung im Signal ausgeschlossen.

5.2.3 F0-Extraktion und Normalisierung

Jede produzierte Silbe wird in Praat (Boersma & Weenink 2009) analysiert. Mit Hilfe eines Praat-Skripts (Xu, 2013) werden die F0-Werte vom stimmhaften Anteil einer annotierten Silbe gemessen und die Werte an zehn gleichabständigen Punkten extrahiert. Durch dieses Verfahren werden F0-Werte eines Tons zeitlich normalisiert, d.h. die Dauer aller tontragenden Silben wird als identisch betrachtet. Anschließend werden die Werte in die folgende Gleichung eingefügt. Die *T-value formula* in (5.3) wandelt die F0-Werte aller Töne eines Sprechers in logarithmische Werte um:

$$(5.3) \quad T = [(lg x - lg min) / (lg max - lg min)] * 5$$

(Shi 1986: 79)

In dieser Formel repräsentiert „x“ die absoluten F0-Werte an einem beliebigen Messpunkt, „min“ steht für das F0-Minimum und „max“ für das F0-Maximum eines Sprechers. Als Ergebnisse der Gleichung erhält man Tonwerte zwischen „1“ und „5“, die der Fünferskala von Chao (1930) entsprechen. Wie in Kapitel 2 erwähnt, erfasst die Fünferskala die phonetischen Eigenschaften eines produzierten Tons und ist aus diesem Grund eine ideale Normalisierung, um verschiedene phonetische Realisierungen eines phonologischen Tons zu vergleichen.

5.2.4 Variabilitätsanalyse

In der vorliegenden Studie hängt die Variabilität in der Tonproduktion von mehreren Faktoren ab. Der erste Faktor sind die **Rahmenbedingungen**, unter denen Töne realisiert werden. Die F0-Kontur eines Tons variiert von der Produktion in Zitierform zu der in Sequenz aufgrund koartikulatorischer Effekte. Der zweite Faktor ist die **Sprechergruppe**. Wenn Töne von Muttersprachler und Lerner produziert werden, weisen die zwei Gruppen an den beiden relevanten Bereichen (Onset und Offset) Variabilität in unterschiedlichem Umfang auf. Der dritte Faktor ist der tonale Kontext, in dem Töne produziert werden. Wenn ein Ton in Sequenz produziert wird, variiert die F0-Kontur je nach Identität der benachbarten Töne. Das heißt, das Onset eines T2 zeigt nach einem T1 weniger Variabilität als nach einem T3. Die drei obengenannten Faktoren werden in die Variabilitätsanalyse der vorliegenden Studie integriert und die Variabilität wird anhand zweier Rahmenbedingungen bei zwei Sprechergruppen und in verschiedenen tonalen Umgebungen ermittelt und verglichen.

Konkret wird die Variabilität in der vorliegenden Studie wie folgt berechnet: Zuerst wird der Mittelwert der normalisierten Tonhöhe (in der Analyse durch *T-value* dargestellt) der ersten und letzten zwei von zehn Messpunkten jeder Silbe in Zitierform berechnet. Weitere Mittelwertberechnungen von *T-value* der ersten zwei Messpunkte werden durchgeführt, wenn die Silbe in der Endposition (P2) eines zweisilbigen Wortes produziert wird, bzw. der letzten zwei Messpunkte, falls die Silbe in der Anfangsposition (P1) produziert wird. Die Auswahl der ersten bzw. letzten zwei von zehn Messpunkten für die Mittelwertberechnung ist ein Kompromiss zwischen der eventuellen Unzuverlässigkeit des einzigen Werts am Onset bzw. Offset und dem Risiko, dass die ermittelte Variabilität einen zu großen Bereich umfasst und möglicherweise neben koartikulatorischen Effekten von weiteren Faktoren beeinflusst wird. Der genaue Wirkungsbereich der beiden koartikulatorischen Effekte wird in Abschnitt 5.3.3 noch einmal aufgegriffen und untersucht.

Der zweite Schritt ist die Ermittlung der Variabilität, die den Standardabweichungen der berechneten Mittelwerte entspricht. Zunächst wird die **Variabilität_{ALLE}** errechnet, die durch die Berechnung der Standardabweichung über drei Wiederholungen pro Zielsilbe, alle Zielsilben mit unterschiedlichen Segmenten¹⁸ und vier tonale Kontexte (T1, T2, T3 und T4) ermittelt wird, ungeachtet der Identität des Tons der Zielsilbe. Zudem wird die **Variabilität_{SEPARAT}** ermittelt, die im Gegensatz zur **Variabilität_{ALLE}** die Standardabweichung separat für jeden der vier Töne auf der Zielsilbe darstellt. Beide Werte, **Variabilität_{ALLE}** und **Variabilität_{SEPARAT}**, werden jeweils von den ersten bzw. letzten zwei Messpunkten der Zielsilben berechnet. Ein Beispiel für die Berechnung der **Variabilität_{SEPARAT}** zeigt Abbildung 5.6.

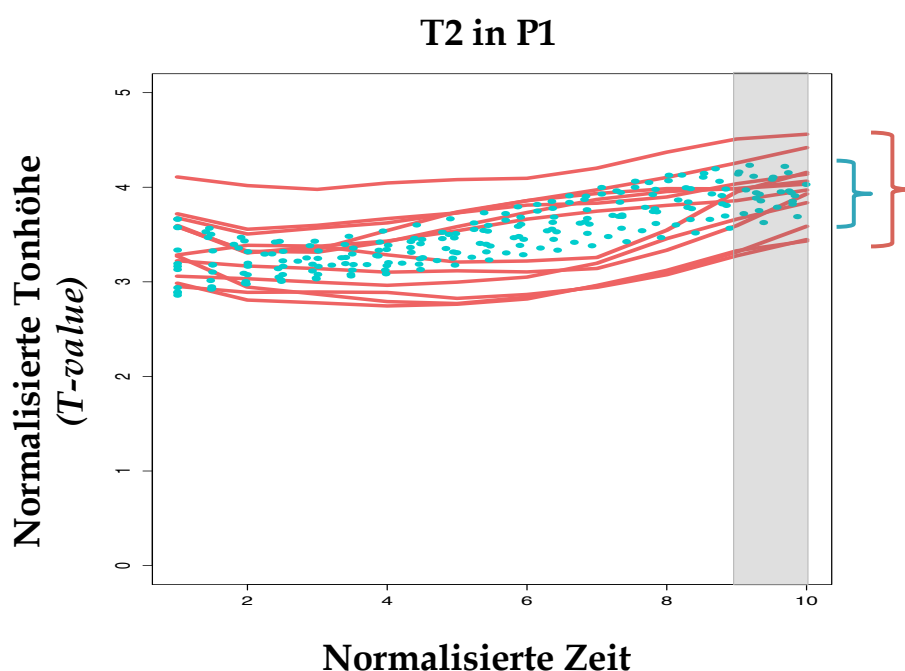


Abbildung 5.6: Darstellung der Berechnung der **Variabilität_{SEPARAT}** am Beispiel von T2 in P1 eines zweisilbigen Wortes. Rote Linien stellen von Lernern und blaue Linien von Muttersprachlern produzierte Konturen dar.

In Abbildung 5.6 sind Konturen eines T2 zu sehen, der in P1 realisiert wird. Die roten Linien repräsentieren die Realisierungen einer der drei Komposita *weimao*, *wumo* und *wami* mit den vier lexikalischen Tönen in P2 durch einen Lerner mit

¹⁸ Wie in Abschnitt 5.2.1 beschrieben, werden für jeden Ton und jede Tonkombination drei Silben mit unterschiedlichen Segmenten eingesetzt.

jeweils drei Wiederholungen, was zwölf Konturen ergibt. Die blauen Linien repräsentieren die entsprechenden Realisierungen eines Muttersprachlers. Auf der Y-Achse sind die normalisierten Tonhöhen mittels *T-Value* dargestellt und auf der X-Achse sind zehn Messpunkte aufgezeichnet. Der grau markierte Bereich stellt den für die Berechnung der Variabilität relevanten Bereich dar. Die geschweiften Klammern visualisieren den Umfang der Variabilität.

Wenn das Offset eines Tons aufgrund der Änderung der Rahmenbedingungen mehr Variabilität aufweist, wenn also mehr Variabilität des Offsets in P1 eines (zweisilbigen) Wortes als in Zitierform auftritt, ist anzunehmen, dass die Richtung der tonalen Koartikulation in diesem Fall regressiv ist. Wenn aber aufgrund der Rahmenbedingungen mehr Variabilität am Onset eines Tons zu finden ist, sich also mehr Variabilität des Onsets in P2 eines (zweisilbigen) Wortes als in Zitierform zeigt, ist anzunehmen, dass die Richtung der tonalen Koartikulation progressiv ist.

Hypothese H1 lässt erwarten, dass die Produktion der Lerner eine signifikant höhere Variabilität am Offset eines Tons in P1 zeigt, als die der Muttersprachler. Beim Onset eines Tons in P2 ist das Verhältnis entsprechend umgekehrt. Das bedeutet, dass die ermittelte **Variabilität_{ALLE}** und **Variabilität_{SEPARAT}** nicht allein von der Veränderung der Rahmenbedingung bestimmt werden, sondern vielmehr von der Interaktion zwischen Rahmenbedingung und Sprechergruppe. Die statistische Modellierung in der vorliegenden Studie basiert auf dieser Annahme (siehe Abschnitt 5.3.2).

5.3 Ergebnis

5.3.1 F0-Konturen

Zuerst werden die F0-Konturen der vier lexikalischen Töne in Zitierform von Muttersprachlern und Lernern graphisch dargestellt (siehe Abbildung 5.7). Um

die Übersichtlichkeit zu gewährleisten, werden an dieser Stelle jeweils nur die Konturen eines Sprechers einer Gruppe dargestellt. In diesem Fall stammen die Konturen von Lerner DL4 (DL = deutsche Lerner) und Muttersprachler CM7 (CM = chinesische Muttersprachler) aus der Produktion der Zielsilben *wumo*. Die vollständigen Kontur-Aufzeichnungen aller Sprecher befinden sich in Appendix D.

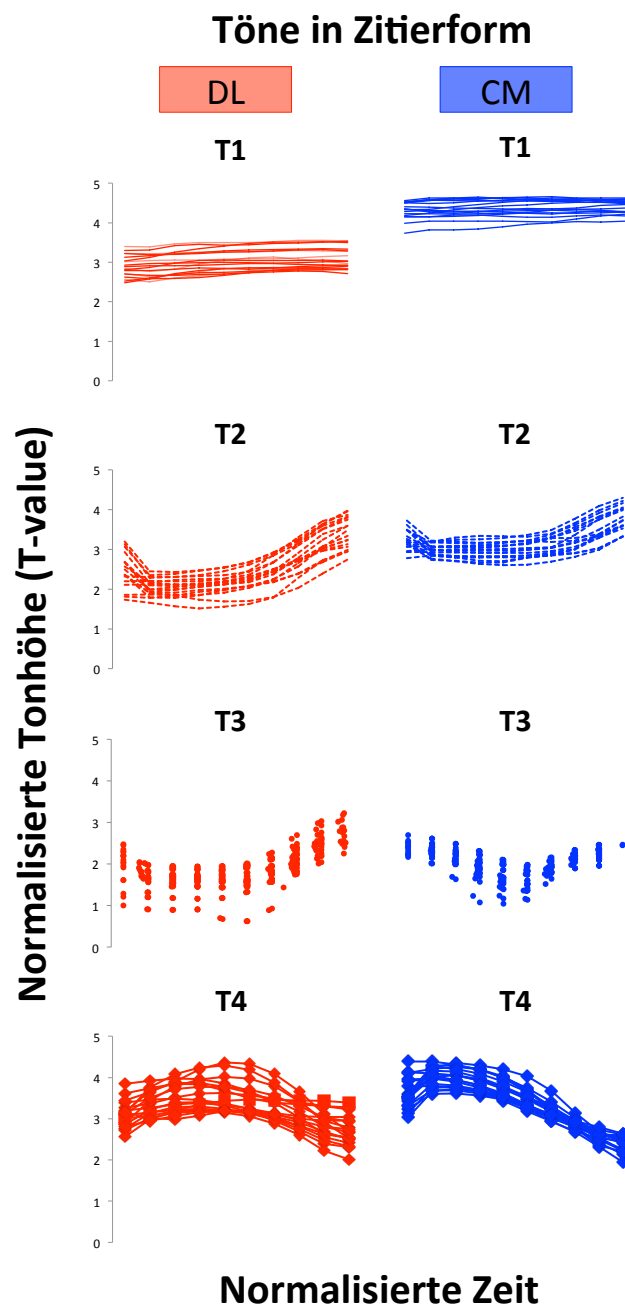


Abbildung 5.7: F0-Konturen von T1, T2, T3 und T4 in Zitierform, produziert von DL4 (linke Spalte) und CM7 (rechte Spalte).

Bei einer rein visuellen Betrachtung kann man nur geringfügige Abweichungen zwischen den Konturen beider Sprecher feststellen. Da sich die Probanden in dieser Studie nicht mehr auf Anfängerniveau des SC befinden und sich explizite Übungen zur Produktion der Töne im Unterricht in der Regel auf die Zitierform beschränkte, entspricht die geringe Abweichung zur muttersprachlichen Produktion den Erwartungen. Dies stimmt auch mit den Ergebnissen in der Literatur überein (z.B.: Yang 2011; He 2010). Insgesamt zeigen die Konturen von DL4 auch innerhalb der eigenen Tonproduktion des jeweiligen Tons wenig Abweichung. Dies könnte darauf hindeuten, dass bei den deutschen Probanden relativ stabile Kategorien von Tönen in Zitierform vorhanden sind.

In den Konturen der Töne zweisilbiger Wörter sind jedoch größere Abweichungen zwischen der Tonproduktion von Lernern und Muttersprachlern festzustellen. In Abbildung 5.8 wird der Umfang der Variabilität am Offset visualisiert. Auf der linken Seite (a.) sind die Konturen von Lernern und auf der rechten Seite (b.) die von Muttersprachlern aufgezeichnet. Die Länge der violetten Pfeile zeigt dabei den Umfang der Variabilität am Offset. In Abbildung 5.9 wird der Umfang der Variabilität am Onset eines produzierten Tons dargestellt. In (c) werden Konturen von Lernern und in (d.) die von Muttersprachlern aufgezeigt. Der Umfang der Variabilität am Onset wird durch die Länge der grünen Pfeile repräsentiert. Jede Zeile der Abbildung besteht aus Konturen eines produzierten Tons in P1 bzw. P2 eines zweisilbigen Wortes.

Zunächst werden die Konturen¹⁹ von in P1 bzw. P2 produzierten Tönen der zwei Sprechergruppen verglichen. In Abbildung 5.8 ist zu erkennen, dass die violetten Pfeile in (a.) immer länger als in (b.) sind. Dies weist darauf hin, dass unter dem Einfluss des folgenden Tons bei den Lernern mehr Variabilität am Offset als bei den Muttersprachlern vorliegt. Umgekehrt sind die grünen Pfeile in Abbildung 5.9 (c.) bei allen vier Tönen kürzer als die in (d.), was darauf hindeutet, dass die

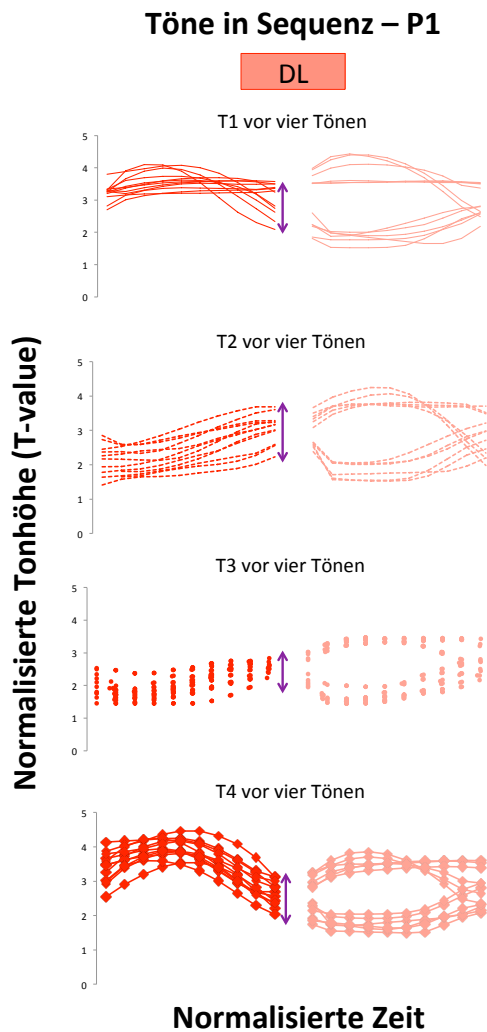
¹⁹ Hier werden Konturen in der Sequenz T3-T3 nicht berücksichtigt. Diese Tonsequenz wird separat in Abschnitt 5.3.5 behandelt.

Realisierung von Tönen in P2 bei Lernern sichtbar weniger von dem vorausgehenden Ton beeinflusst ist als die der Muttersprachler.

Wenn man die Konturen der von Lernern produzierten Töne in P1 (Abbildung 5.8 a.) und in P2 (Abbildung 5.9 c.) vergleicht, stellt man fest, dass die violetten Pfeile stets länger sind als die grünen, da die Variabilität in P1 deutlich größer ist als in P2. Für die Tonproduktion der Muttersprachler gilt ein entgegengesetztes Bild, in dem die grünen Pfeile in Abbildung 5.8 (b.) immer länger sind als die violetten in Abbildung 5.9 (d.). Dies bedeutet, dass die Variabilität in P1 geringer ist als in P2 in der muttersprachlichen Produktion.

Die in Abbildung 5.8 und 5.9 gezeigten Muster weisen darauf hin, dass in der Tonproduktion der Lerner die regressive tonale Koartikulation stärker ist als die progressive. In der Tonproduktion der Muttersprachler dagegen ist die progressive tonale Koartikulation die dominantere, was mit den Ergebnissen in der Fachliteratur übereinstimmt.

(a.)



(b.)

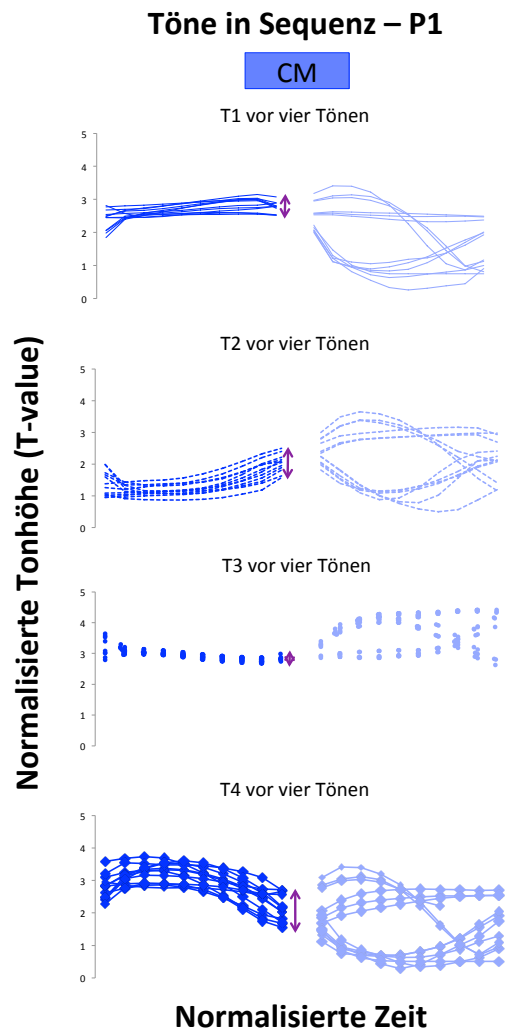
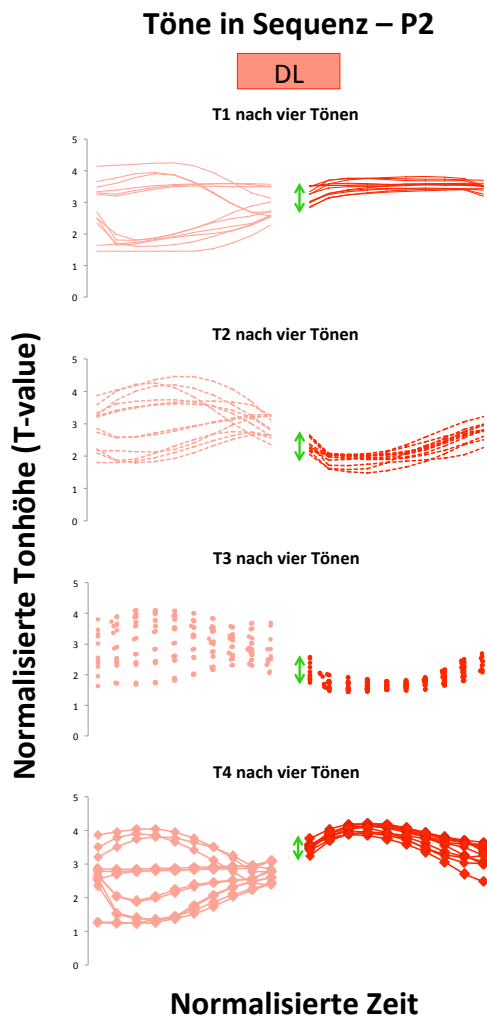


Abbildung 5.8: F0-Konturen von Zielsilben in P1 eines zweisilbigen Wortes. (a.) stellt die Produktion von Lernern und (b.) die Produktion von Muttersprachlern dar. Die einzelnen Zeilen zeigen die Konturen eines Tons gefolgt von vier lexikalischen Tönen. In (a.) und (b.) steht die intensive Farbe für Konturen von Zielsilben, die blasser Farbe für Konturen der vier nachfolgenden Töne.

(c.)



(d.)

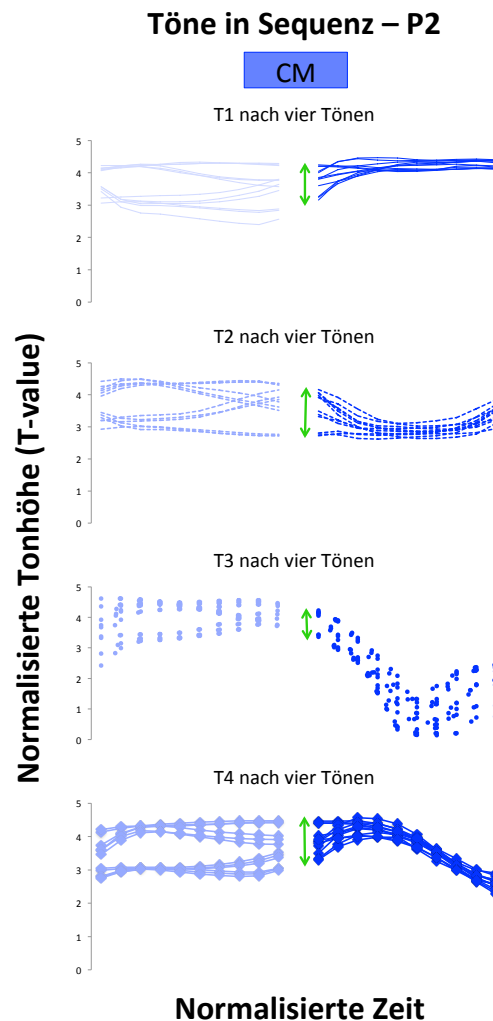


Abbildung 5.9: F0-Konturen von Zielsilben in P2 eines zweisilbigen Wortes. (c.) stellt die Produktion von Lernern und (d.) die Produktion von Muttersprachlern dar. Die einzelnen Zeilen zeigen die Konturen eines Tons, der nach vier Tönen produziert wird. In (c.) und (d.) steht die intensive Farbe für Konturen von Zielsilben, die blasser Farbe für Konturen der vier vorausgehenden Töne.

5.3.2 Statistische Analyse

Wie in Abschnitt 5.2.4 erwähnt, beruht die statistische Modellierung auf der Annahme, dass die Variabilität in den relevanten Bereichen von der Interaktion der Rahmenbedingung (in Zitiertform und in Sequenz) und der unterschiedlichen Sprechergruppe (Lerner und Muttersprachler) bestimmt wird. Entsprechend werden die oben genannten Faktoren innerhalb des statistischen Modells als

CONDITION und GROUP bezeichnet. Sämtliche statistische Analysen der vorliegenden Arbeit werden mithilfe der Statistik-Software *R* durchgeführt.

Für die Untersuchungen zur regressiven und progressiven Koartikulation werden *mixed linear regression models* formuliert, mit der **Variabilität_{ALLE}** bzw. **Variabilität_{SEPARAT}** als abhängige Variablen und der Interaktion von CONDITION und GROUP als *fixed* Faktor. Da in der vorliegenden Studie der Schwerpunkt nicht in der Erforschung der individuellen Variabilität der Sprecher und einzelnen Silbenstrukturen liegt, werden das *random intercept* von jedem Sprecher (Speaker) und jedem segmentalen Inhalt (Item) als die *random* Faktoren eingesetzt.

Zuerst wird die Variabilität von Realisierungen der Zielsilben mit allen vier Tönen zusammengefasst und auf ihre jeweiligen regressiven und progressiven koartikulatorischen Effekte hin untersucht, was zwei *mixed linear regression models* ergibt, wie (5.4) zeigt:

$$(5.4) \quad \text{Variabilität}_{\text{ALLE}} \sim \text{CONDITION} * \text{GROUP} + (1 | \text{Speaker}) + (1 | \text{Item})$$

Likelihood Ratio Tests können eine Signifikanz erreichen, wenn das vollständige Modell mit dem reduzierten Modell ohne die zweifache Interaktion verglichen wird (die **Variabilität_{ALLE}** am Offset für den regressiven Effekt $p < 0.0001$ und die **Variabilität_{ALLE}** am Onset für den progressiven Effekt $p < 0.0001$).

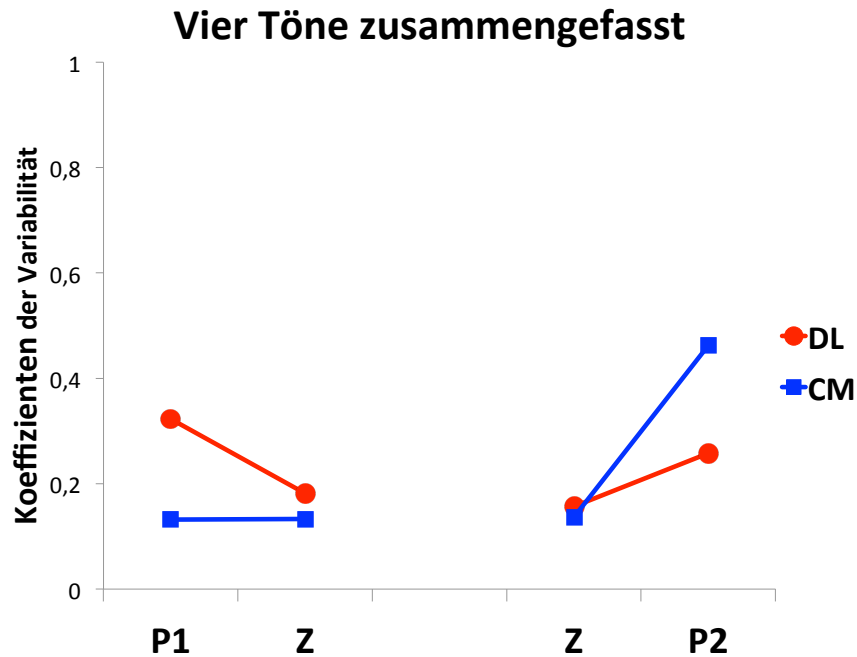


Abbildung 5.10: Koeffizienten der Variabilität im statistischen Modell am Offset (links) und am Onset (rechts) sowie der Interaktion von CONDITION („Z“ = Zitierform vs. „P1“ = Töne produziert in P1 bzw. „P2“ = Töne produziert in P2) und GROUP (Lerner in Rot (DL) vs. Muttersprachler in Blau (CM)).

In Abbildung 5.10 werden die Koeffizienten der Variabilität im obigen statistischen Modell aufgezeichnet. Dabei kann ein generelles Muster festgestellt werden. In der Zitierform gibt es lediglich minimale Unterschiede zwischen den beiden Gruppen. Wenn ein Ton jedoch innerhalb eines Wortes realisiert wird, zeigt die Produktion von Lernern eine eindeutig größere Variabilität als die von Muttersprachlern, wenn der Ton sich in P1 befindet. Wenn der Ton in P2 realisiert wird, zeigt wiederum die Produktion von Muttersprachlern mehr Variabilität als die von Lernern.

In dem obigen Modell wird Variabilität der Realisierung aller vier Töne der Zielsilben zusammengefasst. Ist dieses generelle Muster ebenfalls bei der Realisierung einzelner Töne der Zielsilben zu finden? Verhält sich jeder Ton unter koartikulatorischen Effekten gleich? Um diese Fragen zu beantworten, wird der Datensatz nach den vier lexikalischen Tönen auf der Zielsilbe unterteilt. **Variabilität_{SEPARAT}** wird auf regressiven und progressiven koartikulatorischen Effekt hin untersucht, was acht *mixed linear regression models* ergibt, wie (5.5) zeigt:

(5.5) **Variabilität_{SEPARAT}** \sim **CONDITION*GROUP** + (1 | Speak) + (1 | Item)

Für alle linearen Modelle kann die Signifikanz in den *Likelihood Ratio Tests* erreicht werden. Das Ergebnis wird in der folgenden Tabelle zusammengefasst:

Töne auf der Zielsilbe	Signifikanz der <i>CONDITON*GROUP</i> der Variabilität des Offsets	Signifikanz der <i>CONDITION*GROUP</i> der Variabilität des Onsets
T1	0.01739 *	< 2.2e-16 ***
T2	0.01788 *	2.249e-12 ***
T3	0.003639 **	7.473e-08 ***
T4	0.001004 **	0.001355 **

Tabelle 5.1: Signifikanz der *Likelihood Ratio Tests* für die *mixed linear regression models* der **Variabilität_{SEPARAT}** des Offsets und Onsets.

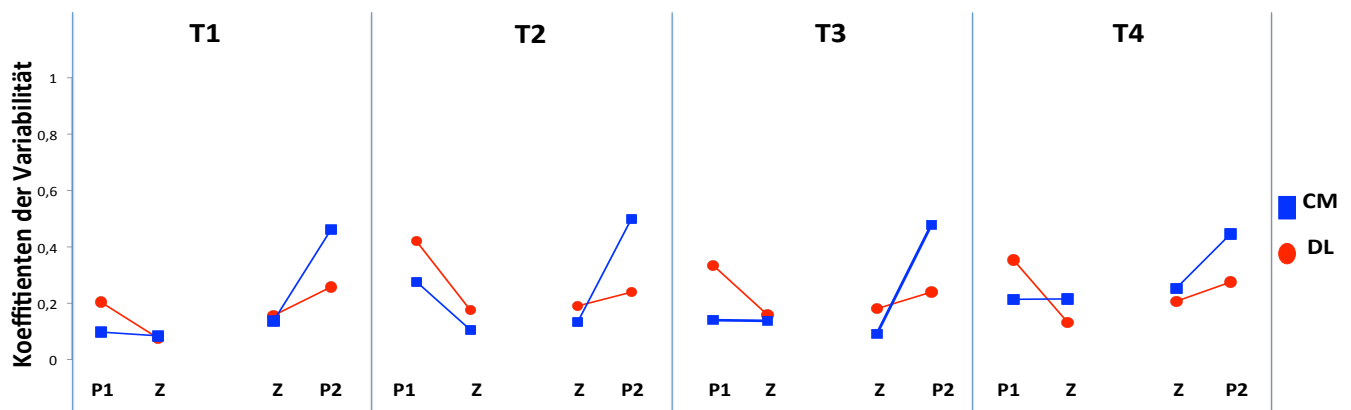


Abbildung 5.11: Die Koeffizienten der Variabilität in statistischen Modellen für einzelne Töne am Offset (links) und am Onset (rechts) in jedem Teilbild sowie der Interaktion von *CONDITION* („Z“ = Zitierform vs. „P1“ = Töne produziert in P1 bzw. „P2“ = Töne produziert in P2) und *GROUP* (Lerner in Rot (DL) vs. Muttersprachler in Blau (CM)).

Auch bei einzelnen Tönen ist deutlich zu sehen, dass sich die Variabilität der tonalen Realisierung im betreffenden Bereich von beider Gruppen in der Zitierform kaum, in einem zweisilbigen Wort jedoch deutlich voneinander unterscheidet. Während bei Lernern die Variabilität des Offsets in der Realisierung aller Töne höher als bei Muttersprachlern ist (rote Linien befinden sich stets oberhalb der blauen Linien in der linken Hälfte jedes Teilbildes), ist bei

Muttersprachlern die Variabilität des Onsets in der Realisierung aller Töne höher als bei Lernern (blaue Linien sind oberhalb der roten Linien). Dieses Muster deutet darauf hin, dass die tonale Realisierung von Lernern bei allen Tönen vorwiegend dem regressiven Effekt und die von Muttersprachlern hauptsächlich dem progressiven Effekt unterliegt.

Dabei sind zwei Phänomene in Abbildung 5.11 bemerkenswert: (1) In der Tonproduktion des T2 in P1 zeigt die Variabilität bei Muttersprachlern ebenfalls eine deutliche Steigerung zur Produktion des T2 in Zitierform. Xu (1997) zeigt in seiner Studie, dass der regressive Effekt im SC sehr schwach ausgeprägt ist und man ihn am deutlichsten beim F0-Maximum einer betroffenen Silbe feststellen kann. Bei einem T2 wird dieser Effekt folglich gegen Ende der Silbe sichtbar, weil das F0-Maximum in der steigenden Kontur von T2 am Silbenende erreicht wird. Auch die Daten der vorliegenden Arbeit können diese Tendenz bestätigen. (2) Die Variabilität am Offset von T4 in Zitierform in der muttersprachlichen Produktion ist sogar höher als in der Lerner-Produktion. In diesem Fall lassen die Muttersprachler augenscheinlich mehr Variabilität zu als die Lerner. Der Grund dafür könnte sein, dass sich bei T4 das Tontarget oder die maximale Annäherung daran nicht am Offset befindet, sondern früher, was Xus & Wangs Implementierungsregel (2001) widerspricht, die besagt, dass Targets der Töne im SC gegen Ende der tontragenden Silbe erreicht werden. Tatsächlich stellt der Autor der vorliegenden Arbeit nach längerer Unterrichtserfahrung im Bereich ChaF fest, dass deutsche Sprachschüler T4 erst korrekt aussprechen, wenn sie den Startpunkt (Wendepunkt) der fallenden Tonbewegung von T4 treffen. Beides deutet darauf hin, dass die Tontargets im SC nicht ausschließlich binär beschrieben werden sollen. Für eine korrekte Produktion des T4 ist sowohl das Register als auch die F0-Bewegung entscheidend.

5.3.3 Wirkungsbereich der koartikulatorischen Effekte

Im vorangegangenen Abschnitt wird gezeigt, dass die Richtung der tonalen Koartikulation in der Tonproduktion von Lernern und Muttersprachlern

entgegengesetzt ist. Jedoch ist die Richtung nicht der einzige Faktor, den man in der Untersuchung der tonalen Koartikulation berücksichtigen muss. Die nächste Frage bezieht sich auf den Wirkungsbereich der beiden koartikulatorischen Effekte. Xu (1997) berichtet, dass der progressive Effekt am Onset der betroffenen Silbe deutlich zu sehen ist und er bis zur Hälfte der Silbe reichen kann. Der regressive Effekt dagegen beeinflusst, falls vorhanden, das F0-Maximum der betroffenen Silbe (Xu 1997). Brunelle (2009) stellt allerdings fest, dass die zwei unterschiedlichen Richtungen der tonalen Koartikulation auch zwei unterschiedliche Wirkungsbereiche haben. Er entdeckt in seinen Daten aus dem Vietnamesischen, dass der regressive Effekt schwach, aber dennoch langfristig wirkend ist. Der Einfluss des regressiven Effekts kann bis zum Onset der betroffenen Silbe reichen und bewegt die gesamte Höhe der Kontur als Ganzes auf- bzw. abwärts. Der progressive Effekt hat einen lokalen Einfluss, der stark, jedoch innerhalb kurzer Zeit abnehmend ist (Brunelle 2009). Dass die Richtung der tonalen Koartikulation in der Tonproduktion bei Lernern und Muttersprachlern einander entgegengesetzt ist, gibt keinen Aufschluss darüber, ob sich der Wirkungsbereich der tonalen Koartikulation bei der Realisierung der Lerner und Muttersprachler ebenfalls unterscheidet.

Um den Wirkungsbereich der beiden koartikulatorischen Effekte zu ermitteln und zwischen zwei Sprechergruppen zu vergleichen, werden die gesamten Daten nach Gruppen aufgeteilt. In jedem Teildatensatz wird statt der **Variabilität_{SEPARAT}** der ersten bzw. letzten zwei Punkte (wie in Abschnitt 5.2.4) die aller zehn Messpunkte einer Silbe berechnet. Diese **Variabilität_{SEPARAT}** wird in einem *mixed linear regression model* als abhängige Variable eingesetzt, alleine die CONDITION ist der *fixed* Faktor, das *intercept* des Sprechers und der segmentale Inhalt sind die *random* Faktoren. Das bedeutet, dass für jede Gruppe und koartikulatorische Richtung zehn *mixed models* gebildet werden. Die zehn Modelle für jede koartikulatorische Richtung werden alle in *Likelihood Ratio Tests* mit reduzierten Modellen ohne den Faktor CONDITION getestet. Anschließend wird der *Chisq* Wert aus den *Likelihood Ratio Tests* erfasst, um die Einflussstärke des *fixed* Faktors CONDITION darzustellen. Je höher der *Chisq* Wert, desto stärker beeinflusst dieser Faktor die

abhängige Variable. Das statistische Modell ist in (5.6) dargestellt und die Ergebnisse werden in Abbildung 5.12 dargestellt:

$$(5.6) \quad \text{Variabilität}_{\text{SEPARAT}} \sim \text{CONDITION} + (1 \mid \text{Speak}) + (1 \mid \text{Item})$$

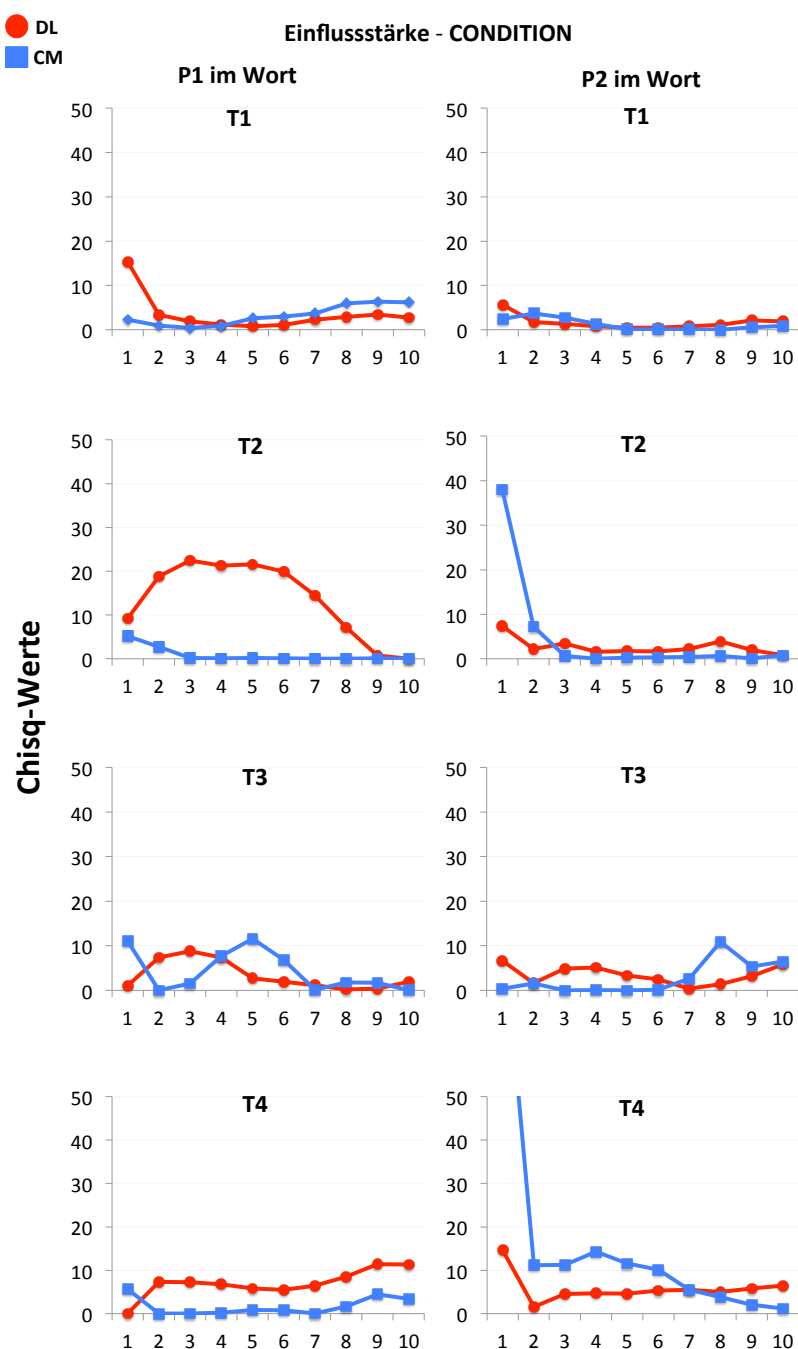


Abbildung 5.12: Wirkungskraft der Variable **CONDITION** auf die Variabilität der Messpunkte von Tönen in P1 (linke Spalte) und P2 (rechte Spalte) bei Lernern (rot) und Muttersprachlern (blau).

Durch Veränderungen in der Kurve in Abbildung 5.12 (rot für Lerner und blau für Muttersprachler) wird visualisiert, inwiefern der Einfluss des koartikulatorischen Effekts reicht und ob sich der Wirkungsbereich zwischen Lernern und Muttersprachlern unterscheidet. Alle Linien in der linken Spalte zeigen den Wirkungsbereich des regressiven Effekts. Die roten Linien zeigen die starke Tendenz, dass die Erhebung der Kurve sich nicht auf eine einzelne Stelle am Offset beschränkt, sondern über beinahe alle Messpunkte erstreckt - besonders bei T2 und T4 - während die blauen Linien insgesamt wenig Veränderung zeigen.

Die Linien in der rechten Spalte stellen den Wirkungsbereich des progressiven Effekts dar. Hier zeigen wiederum die blauen Linien die Tendenz, dass die Erhebung der Kurve sich mehr oder weniger auf die ersten Messpunkte beschränkt - insbesondere bei T2 und T4.

Zusammenfassend zeigt sich die Einflussstärke der CONDITION (d.h., dass ein Ton in Zitierform und im Wort produziert wird) am deutlichsten bei der Realisierung von T2 und T4. Dabei wird der Unterschied zwischen den Wirkungsbereichen des regressiven und progressiven Effekts klar. Der progressive Effekt wirkt kurzfristig und stark, was sich bei der Produktion von Tönen in P2 durch Muttersprachler feststellen lässt. Der regressive Effekt wirkt langfristig und schwach, was sich besonders in der Tonproduktion von T2 und T4 bei Lernern zeigt. Das heißt, nicht nur Muttersprachler einer Tonsprache wie etwa Vietnamesisch (Brunelle 2009), sondern auch Lerner einer Nicht-Tonsprache haben einen ähnlichen Wirkungsbereich des regressiven Effekts. Der Wirkungsbereich tonaler koartikulatorischer Effekte scheint sprachübergreifend zu sein.

5.3.4 Assimilation vs. Dissimilation

Um die Hypothese H2 zu überprüfen, wird das Wesen der jeweiligen koartikulatorischen Effekte untersucht. Es wird über tonale Koartikulation im SC berichtet, dass der progressive Effekt eindeutig assimilatorisch und der regressive

Effekt (falls vorhanden) dissimilatorisch ist. Wird ein T1 nach einem T3 produziert, ist der Onset-Wert des T1 niedriger als in der Zitierform. Der Onset-Wert des T1 wird durch den tiefen Offset-Wert des T3 assimiliert. In Abschnitt 5.3.2 wird festgestellt, dass Lerner in der Tonproduktion einen stärkeren regressiven Effekt aufweisen. Ob dieser Effekt assimilatorischer oder dissimilatorischer Natur ist, wird im Folgenden untersucht. Dafür wird der Mittelwert (statt der Variabilität) von *T-value* der letzten zwei Messpunkte von Tönen in P1 und der ersten zwei Messpunkte von Tönen in P2 ermittelt.

In Abbildung 5.13 wird der regressive Effekt auf den Mittelwert der *T-values* der letzten zwei Messpunkte einer tonalen Realisierung dargestellt. Auf der linken Seite sind die Werte der Lerner zu sehen. Das gelbe Kästchen zeigt dabei die Werte der Töne in Zitierform und die grünen Kästchen die Werte des jeweiligen Tons vor einem T1, T2, T3 und T4. Auf der rechten Seite sind die Werte der Muttersprachler zusammengefasst. Hier zeigt das pinkfarbene Kästchen auf der linken Seite den Wert eines Tons in Zitierform und die violetten Kästchen die Werte des jeweils Tons vor einem T1, T2, T3 und T4. Auf der X-Achse ist die Rahmenbedingung gekennzeichnet, in der ein Ton produziert wird. Ein Ton wird entweder in Zitierform („z“) oder gefolgt von einem der vier Tönen („T1“ „T2“ „T3“ und „T4“) produziert. Auf der Y-Achse sind die Mittelwerte in form von *T-values* der letzten zwei Messpunkten der Zielsilbe dargestellt.

In Abbildung 5.13 kann man einen regressiven Effekt in der Tonproduktion der Lerner (DL) am deutlichsten bei T2 und T3 beobachten. Die Offset-Werte von T2 und T3 sinken sichtbar aufgrund der hohen Onset-Werte der nachfolgenden T4 und T1 ab. Das deutet darauf hin, dass in der L2-Tonproduktion der regressive Effekt von dissimilatorischer Natur ist. In der Tonproduktion der Muttersprachler ist der regressive Effekt am deutlichsten bei T2 gefolgt von T3 (durch einen roten Kreis gekennzeichnet). In diesem Fall ist der Offset-Wert des T2 aufgrund des niedrigen Onset-Werts des nachfolgenden T3 gestiegen, was ebenfalls eine Dissimilation ist.

Regressiver Effekt

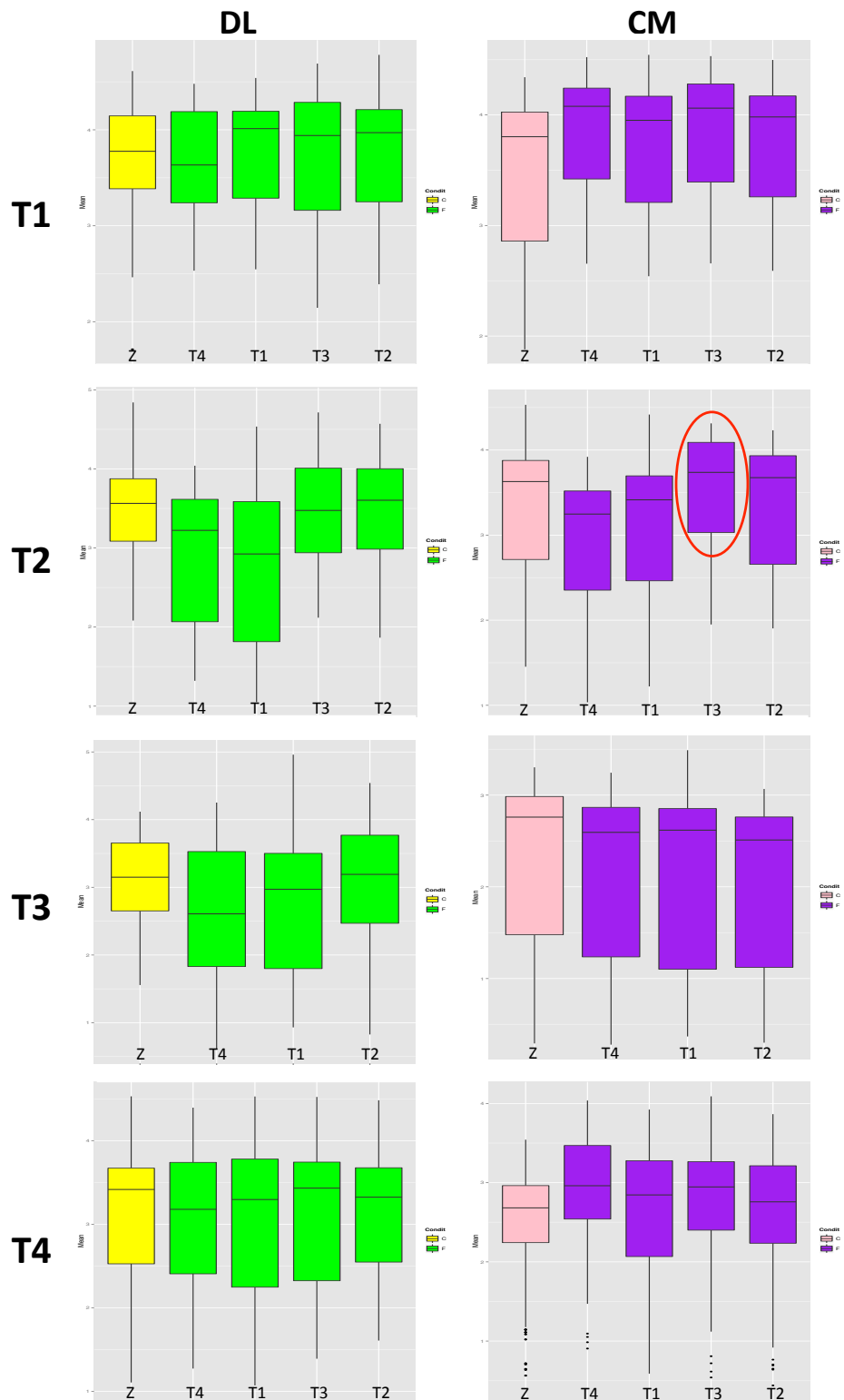


Abbildung 5.13: Regressiver Effekt auf den Mittelwert der letzten zwei Messpunkte einer Silbe in P1. Die linke Spalte stellt die Lerner und die rechte Spalte die Muttersprachler dar. „z“ steht für Mittelwerte der *T-values* der Tonproduktion in Zitierform, „T1“ „T2“ „T3“ und „T4“ steht für Mittelwerte der Tonproduktion in P1.

Abbildung 5.14 stellt den progressiven Effekt auf die Mittelwerte der *T-values* der ersten zwei Punkte der tonalen Realisierung dar. Die Farbkodierung ist mit der Abbildung 5.13 identisch. Auf der X-Achse wird die Rahmenbedingung dargestellt, in der ein Ton produziert wird. Er wird entweder in Zitierform („z“) oder hinter einem der vier Töne („T1“ „T2“ „T3“ und „T4“) produziert. Auf der Y-Achse sind die Mittelwerte in form von *T-values* der ersten zwei Punkte der Zielsilbe dargestellt.

Die Tonproduktion der Muttersprachler in Abbildung 5.14 zeigt ein konstant und systematisch assimilatorisches Muster. Die Onset-Werte der Töne sind nach T3 und T4 mit ihren niedrigen Offset-Werten tiefer als jene in Zitierform. Die Onset-Werte sind mit Ausnahme der Tonsequenz T3-T3 (rot eingekreist) nach den hohen Offset-Werten der T2 und T1 höher. Im Fall der Tonsequenz T3-T3 ist der Onset-Wert des T3 auch höher, obwohl er nach einem Ton mit einem niedrigen Offset-Wert auftritt, nämlich einem T3. Dies ist dadurch zu erklären, dass sich aufgrund des T3-Sandhi die Form des ersten T3 zu einem T2 wandelt, bei dem die phonetische Realisierung einen hohen Offset-Wert hat. Die Tonproduktion von Lernern zeigt dagegen in Abbildung 5.14 nur einen geringfügigen progressiven Effekt, der ebenfalls von assimilatorischer Natur ist. Dieses Ergebnis bestätigt die Hypothese H2, dass Lerner bei der Produktion von Zielwörtern in Zitierform hauptsächlich eine dissimilatorische Tendenz zeigen.

Progressiver Effekt

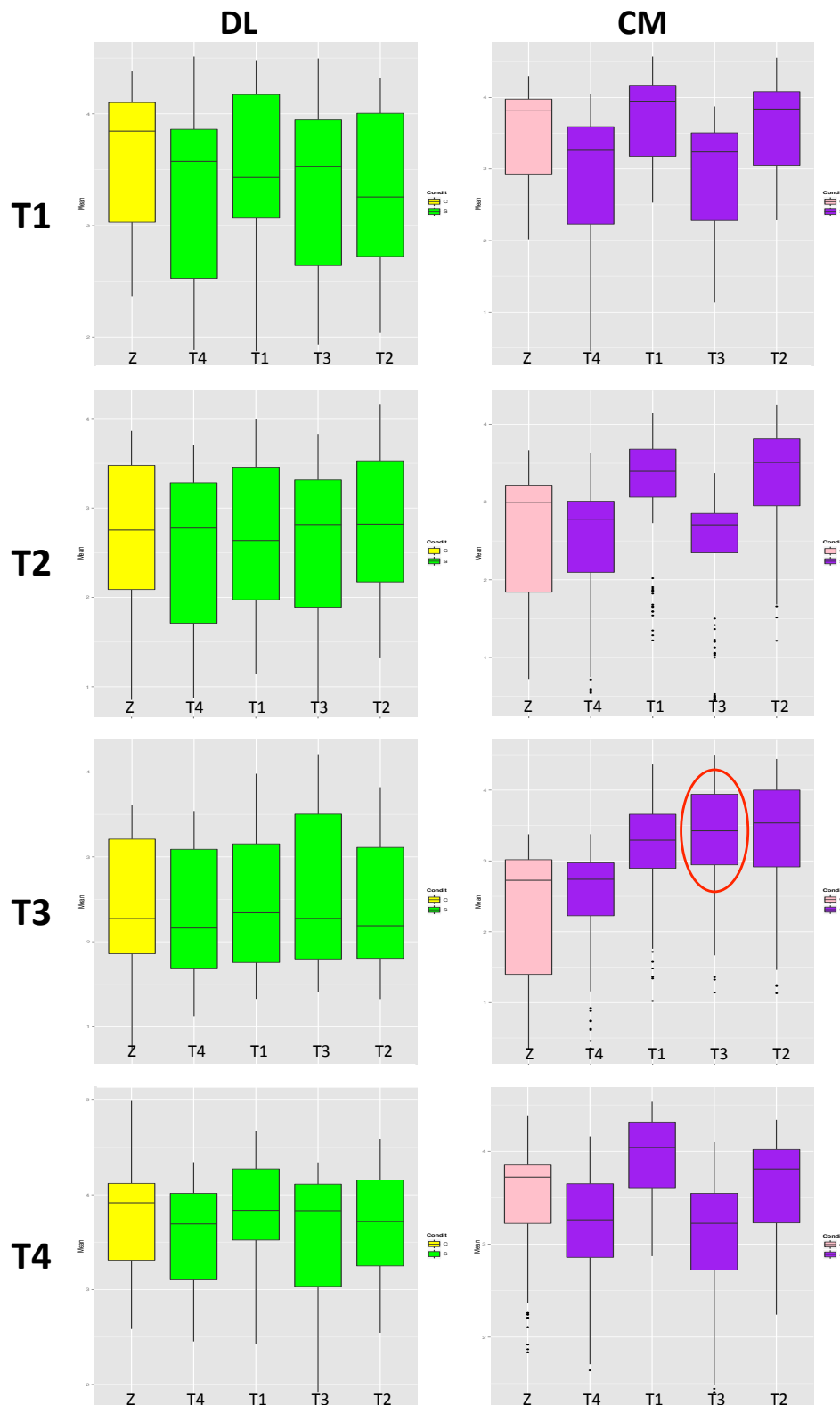


Abbildung 5.14: Progressiver Effekt auf den Mittelwert der ersten zwei Messpunkte von Silben in P2. Die linke Spalte zeigt die Lerner und die rechte Spalte die Muttersprachler. „Z“ steht für Mittelwerte der *T-values* der Tonproduktion in Zitierform, „T1“ „T2“ „T3“ und „T4“ steht für Mittelwerte der *T-values* der Tonproduktion in P2.

Zusammengefasst ist eine regressive dissimilatorische tonale Koartikulation vorwiegend in der Lerner-Produktion zu finden, während die Muttersprachler in der tonalen Koartikulation hauptsächlich progressive Assimilation zeigen.

5.3.5 Tonale Realisierung der Tonsequenz T3-T3

In der obigen Analyse ist die Realisierung der Tonsequenz T3-T3 nicht enthalten, weil sich bei der phonetischen Realisierung aufgrund der Tonsandhi-Regel die tonale Kategorie ändert. Diese Tonsandhi-Regel ist eine der wenigen, die im Lehrmaterial und Unterricht explizit beschrieben und geübt wird. Die Realisierung der Tonsequenz T3-T3 von Lernern und Muttersprachlern sind in Abbildung 5.15 aufgezeichnet.

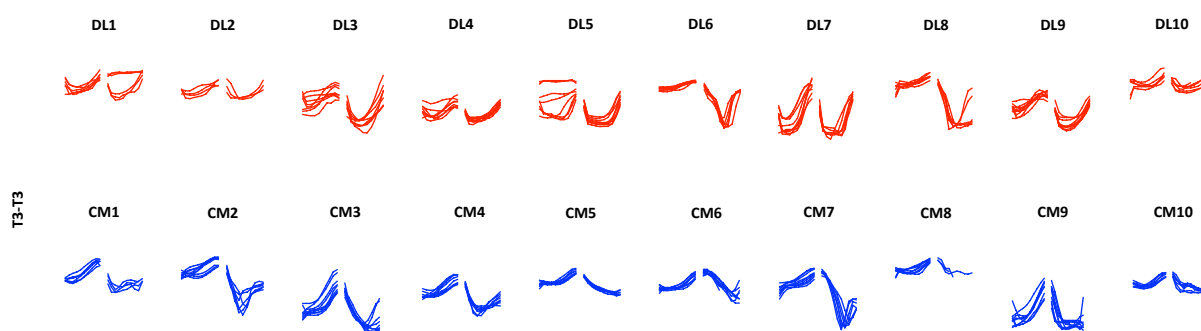


Abbildung 5.15: Tonale Realisierung der T3-T3 Sequenz mit unterschiedlichem segmentalem Inhalt und drei Wiederholungen von Lernern (rot) und Muttersprachlern (blau).

In der Abbildung 5.15 ist zu erkennen, dass bis auf wenigen Realisierungen von DL1 und DL5, alle Lerner in der Lage sind, diese Tonsequenz mit muttersprachlichen Konturen zu produzieren.

Um zu überprüfen, ob die von Lernern und Muttersprachlern produzierten Konturen sich voneinander unterscheiden, wird der Wert der *T-value* von jedem der zehn gemessenen Punkte als abhängige Variable in ein *mixed linear regression model* eingesetzt, wobei die Sprechergruppe (Lerner vs. Muttersprachler) der *fixed* Faktor und *intercept* des Sprechers der *random* Faktor ist. Das statistische Modell wird wie in (5.7) gebildet:

$$(5.7) \quad F0_Wert \sim GROUP + (1 | Speak)$$

(von jedem Messpunkt)

Mittels der Funktion *effect ()* werden aus den Ergebnissen der 20 einzelnen *mixed linear regression models* die obere 95%- und untere 5%-Vertrauensgrenze für jeden Messpunkt ermittelt. Durch die obere und untere Vertrauensgrenze kann man eine Zone festlegen, in der die Messpunkte der Realisierungen beider Gruppen höchstwahrscheinlich liegen (siehe Abbildung 5.16). Demnach ist ein bestimmtes Muster vorauszusagen, das Lerner oder Muttersprachler bei der Produktion einer bestimmten Tonsequenz am wahrscheinlichsten zeigen.

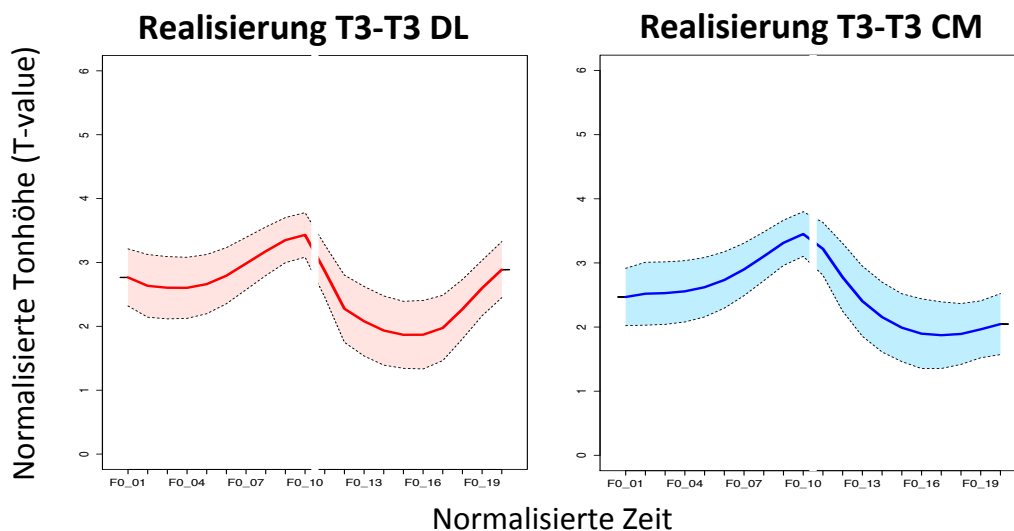


Abbildung 5.16: Obere und untere Vertrauensgrenze der Realisierung von Tonsequenz T3-T3 in Zitierform von Lernern (links) und Muttersprachlern (rechts).

Abbildung 5.16 zeigt ein äußerst ähnliches Muster bei der Realisierung der Tonsequenz T3-T3 bei Lernern und Muttersprachlern. Die Ergebnisse der 20 *mixed linear regression models* ergeben keinen signifikanten Unterschied bei den jeweiligen Messpunkten in den produzierten Konturen. In Tabelle 5.2 sind die Ergebnisse aufgelistet.

Messpunkte	Signifikanz des Unterschieds zwischen DL und CM
F0_01	0,359
F0_02	0,7462
F0_03	0,8302
F0_04	0,8938
F0_05	0,9043
F0_06	0,8518
F0_07	0,7747
F0_08	0,7904
F0_09	0,8855
F0_10	0,9392
F0_11	0,2382
F0_12	0,1943
F0_13	0,4108
F0_14	0,5708
F0_15	0,7442
F0_16	0,9423
F0_17	0,7837
F0_18	0,2722
F0_19	0,0558
F0_20	0,01869*

Tabelle 5.2: Ergebnisse der *mixed linear regression models* beim Vergleich der Werte der Messpunkte von Lernern und Muttersprachlern bei der Produktion der Tonsequenz T3-T3

Dieses Ergebnis stützt die Hypothese H3, dass explizit gelernte Tonsequenzen zweisilbiger Wörter von deutschen ChaF-Lernern ähnlich wie von chinesischen Muttersprachlern realisiert werden.

5.4 Tonale Realisierung der Zielwörter in Äußerungen

In der vorliegenden Studie werden alle Zielwörter neben der Zitierform auch in Äußerungen realisiert, die in einen kurzen Dialog (siehe Appendix C) eingebaut sind. Die Tonproduktion der Zielwörter in Äußerungen durch Lerner und Muttersprachler zeigt kein von der Produktion der Zielwörter in Zitierform abweichendes Muster, nämlich vorwiegende regressive Dissimilation in der Lerner-Produktion und progressive Assimilation in der muttersprachlichen Produktion. Über tonale Koartikulation im SC wird von Xu (1997) berichtet, dass sich die zugrundeliegenden Muster tonaler Koartikulation bei zweisilbigen Wörtern in Zitierform und Äußerungen nicht unterscheiden:

Although the tones of /ma ma/ sequences produced both in isolation and in carrier sentences were analyzed , only the latter are presented here , largely because essentially the same patterns were found for both conditions , differing only in magnitude.

[Obwohl die Tonproduktion der Sequenzen /ma ma/ in Zitierform und im Trägersätzen analysiert wird, werden hier nur Letztere präsentiert, hauptsächlich weil im Wesentlichen dieselben Muster für beiden Bedingungen gefunden wurden, die sich nur in ihrem Ausmaß unterscheiden.]

(Xu 1997: 68, meine Übersetzung)

Die in der vorliegenden Studie gefundenen Muster stimmen bezüglich tonaler Koartikulation bei Wörtern in Zitierform und Äußerungen mit dem Ergebnis von Xu überein. In den Daten der vorliegenden Studie fallen zahlreichen Konturfehler bei der Tonproduktion der Zielwörter in Äußerungen bei einigen Lernern auf. Zum Beispiel zeigt DL4 bei der Produktion von T4-T3 und DL5 bei der Sequenz T2-T4 mit den Zielsilben *wami* in Zitierform und in Äußerungen eindeutig unterschiedliche Konturen, wie Abbildung 5.17 darstellt:

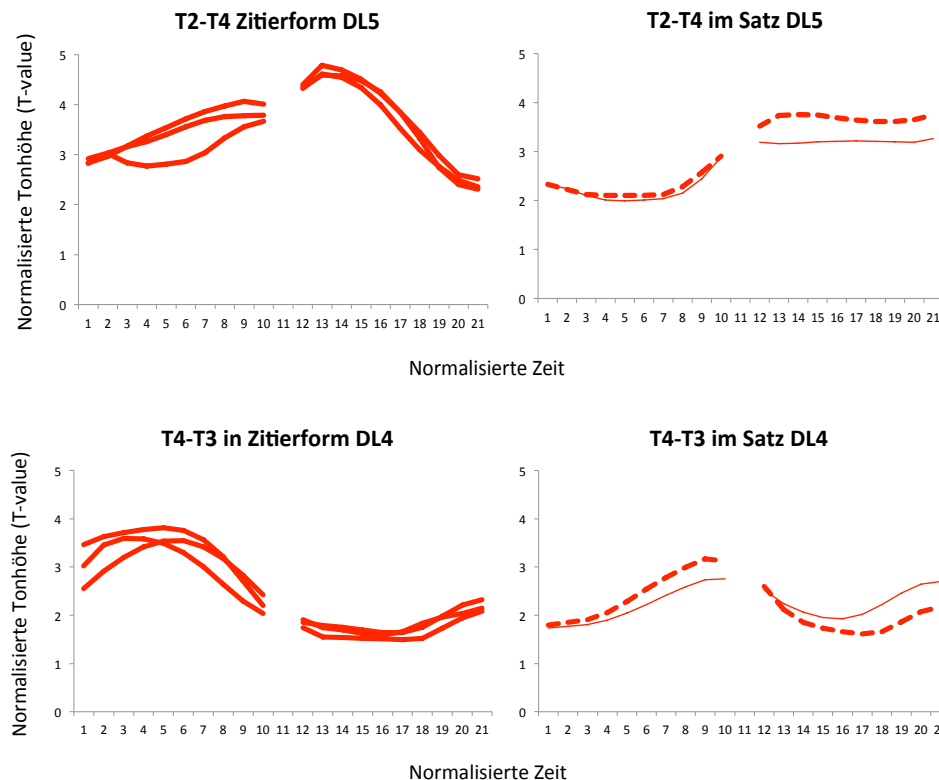


Abbildung 5.17: Beispiel für Konturfehler in der Produktion von Zielwörtern in Zitierform und in Äußerungen bei DL4 und DL5.

Derartige Konturfehler in Abbildung 5.17 könnten ebenfalls ein Ergebnis der regressiven Dissimilation sein. Die typische Auswirkung einer regressiven Dissimilation bei tonaler Koartikulation ist beispielsweise das Anheben eines Tons aufgrund des niedrigen Onset-Wertes des nachfolgenden Tons. In der oberen Reihe hat der hohe Onset-Wert des T4 ungefähr zwei Drittel der Kontur von T2 zum Absinken gebracht, sodass die steil steigende Kontur von T2 zu einer viel flacheren wird. In der unteren Reihe von Abbildung 5.17 hebt der niedrige Onset-Wert des T3 den voraussendenden T4 an, sodass dessen kanonisch fallende Kontur (auf dem linken Bild) zu einer steigenden Kontur (auf dem rechten Bild) verändert wird. Die Konturfehler treten bei Lernern meistens in der Produktion von Äußerungen auf. Dies deutet darauf hin, dass die tonale Realisierung von Tonsequenzen in einer größeren Domäne eine größere Herausforderung für Lerner darstellt.

Bemerkenswert bei der Produktion von Zielwörtern in Äußerungen ist außerdem, dass einige Lerner bei der Tonproduktion auffällig lange Pausen zwischen der ersten und zweiten Silbe einlegen. Abbildung 5.18 zeigt ein paar Beispiele.

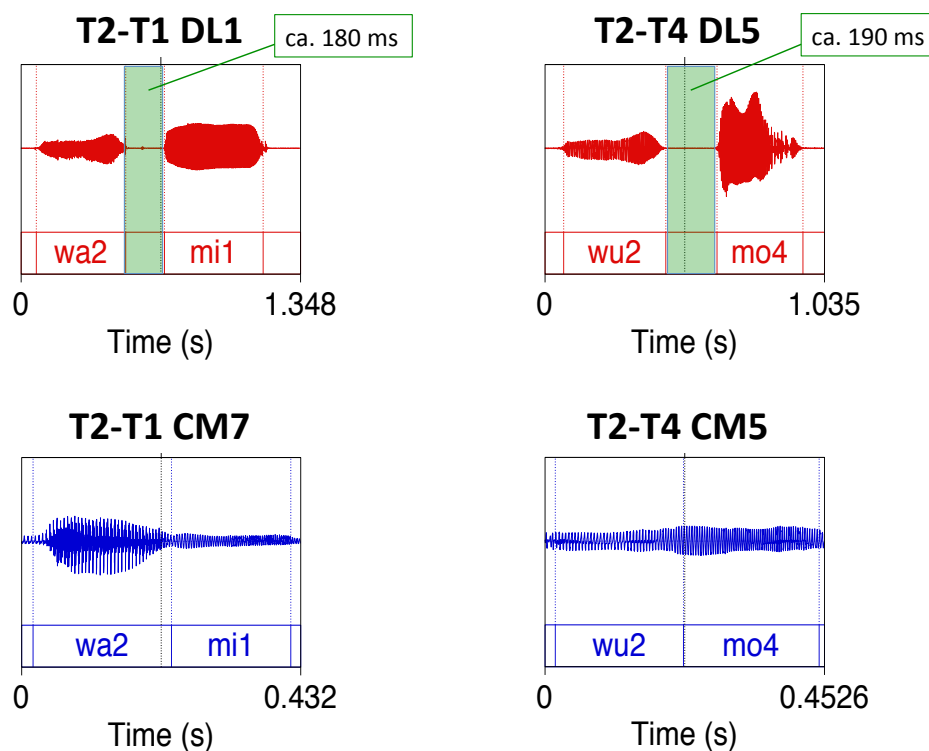


Abbildung 5.18: Beispiele für stark verlängerte Pausen zwischen zwei Silben in den Nonsens-Zielwörtern „wámí“ (linke Spalte) und „wúmò“ (rechte Spalte) bei den deutschen Lernern DL1 und DL5 (obere Zeile, rot gezeichnet). Zum Vergleich werden Tonproduktionen der gleichen Zielwörter bei den Muttersprachlern CM5 und CM7 dargestellt (untere Zeile, blau gezeichnet).

Dieses Phänomen stützt die Hypothese, dass Lerner aufgrund fehlender Übung versuchen, Töne in Sequenz mit der gelernten Zitierform zu realisieren.

5.5 Zusammenfassung der Ergebnisse

In diesem Kapitel wird die Tonproduktion von ein- oder zweisilbigen Wörtern in Zitierform und in Äußerungen bei Lernern und Muttersprachlern untersucht. Im Ergebnis ist zu erkennen, dass die Richtung tonaler Koartikulation in der Tonproduktion der Lerner und Muttersprachler entgegengesetzt ist: Bei Lernern findet sich überwiegend ein regressiver und bei Muttersprachlern ein progressiver Effekt. Dieses Ergebnis bestätigt die aufgestellte Hypothese H1. Der überwiegend

regressive Effekt in der Produktion der Lerner ist von dissimilatorischer Natur, was Hypothese H2 stützt, dass Lerner bei der Tonproduktion versuchen, Töne maximal voneinander abzugrenzen und aus diesem Grund Töne dissimilieren. Bei der Untersuchung zur Tonproduktion der im Unterricht explizit geübten Tonsequenz T3-T3 zeigen Lerner und Muttersprachler bei der tonalen Realisierung sehr ähnliche Konturen. Ein Ergebnis, das Hypothese H3 stützt. Trotz der hier dargestellten Haupttendenz in der Tonproduktion, zeigen deutsche Lerner insbesondere in der Tonproduktion der Zielwörter in Äußerungen auch viel individuelle Variation. Dieses Ergebnis unterstützt die *Interlanguage Hypothesis*, das die Lernersprache ein individuelles Sprachsystem repräsentiert.

6 Realisierung von Wörtern mit Neutralton

Wie in Kapitel 2 ausführlich beschrieben, wird der Neutralton bereits in einigen Studien untersucht. Allerdings behandelt keine der Studien die tonale Realisierung des Neutraltons in der L2-Produktion. Bisher stehen nur die vier vollen lexikalischen Töne im SC im Fokus der Forschung zum Tonerwerb bei Nicht-Tonsprachlern. Die vorliegende Studie versucht, diese Lücke zu schließen und untersucht die tonale Realisierung von Wörtern mit einem Neutralton²⁰ sowohl in Zitierform als auch in Äußerungen von deutschen ChaF-Lernern.

6.1 Forschungshypothese

6.1.1 Metrische Struktur

In der Literatur wird viel über die Existenz des Wortakzents als solchen im SC diskutiert. Dabei sind sich die meisten Autoren bezüglich der metrischen Struktur der zweisilbigen Wörter mit einem Neutralton darin einig, dass sie Trochäen bilden. Dabei ist die Silbe mit einem vollen Ton metrisch stark und die mit dem Neutralton schwach. Duanmu (2004) stellt eine Ähnlichkeit zweisilbiger Wörter mit Neutralton im SC und englischer trochäischer Wörter fest. Im Deutschen sind derartige Trochäen üblich und haben wie im Englischen ebenfalls eine klare Zuweisung des Wortakzentes: die vorletzte Silbe erhält den Wortakzent (z.B. Giegerich 1985; Vennemann 1990) und die letzte Silbe ist unbetont.

Modelle des Zweit- bzw. Fremdspracherwerbs, wie *The Native Language Magnet Model* von Kuhl (Kuhl 1992; Kuhl & Iverson 1995) besagen, dass eine phonetische Einheit in der Zielsprache, die einer Kategorie in der Muttersprache sehr ähnlich jedoch nicht identisch ist, besonders schwer zu erwerben ist. Mit anderen Worten führt eine ähnliche Struktur in der Ziel- und Muttersprache bei Lernern leicht zu

²⁰ Der in der vorliegenden Arbeit behandelte Neutralton beschränkt sich auf den intrinsischen Neutralton nach Shen (1992).

Verwechslungen. Auch *The Speech Learning Model* (Flege 1995) behauptet, dass ähnliche Laute schwierig zu erlernen seien, weil Lerner sie mit denen in der Muttersprache gleichsetzen. Obwohl diese Modelle keine expliziten Vorhersagen über den Erwerb im prosodischen Bereich treffen, ist zu erwarten, dass die ähnliche metrische Struktur in zweisilbigen Wörtern mit einem Neutralton im SC und einem Trochäus im Deutschen zu Schwierigkeiten in der tonalen Produktion führen.

6.1.2 Tonale Realisierung

Ist eine ähnliche metrische Struktur gegeben, ergibt sich daraus nicht zwangsläufig eine ähnliche tonale Realisierung bei zweisilbigen Wörtern mit einem Neutralton im SC und Trochäen im Deutschen. Es existieren zahlreiche Untersuchungen zur Oberflächenkontur des Neutraltons. Shih (1987) listet die F0-Verläufe eines Neutraltons nach den vier vollen lexikalischen Tönen wie folgt auf:

Vorausgehende Töne		Tonlose Silben
T1	H	startet hoch und fällt dann
T2	LH	startet hoch und fällt dann, jedoch nicht so tief wie nach „H“
T3	L	startet relativ tief und steigt dann
T4	HL	startet relativ tief und fällt noch tiefer

Tabelle 6.1: Oberflächenkontur des Neutraltons nach T1, T2, T3 und T4, adaptiert von Shih (1987)

Shih (1987) behauptet, dass sich die F0-Kontur eines Neutraltons in der Satzmitte und in einem bestimmten tonalen Kontext am besten mit einer Interpolationslinie zwischen zwei Tontargets erklären lässt. Eine ähnliche Interpolationstheorie vertreten auch van Santen et al. (1998) und Li (2003). Andere Autoren wie Yip (1980) erklären die F0-Kontur des Neutraltons mit Hilfe der Tonausbreitung (*tonal spreading*). Die steigende Kontur nach T3 sowie die weiter fallende Kontur nach T4 sind dadurch zu erklären, dass sich die final steigende Kontur von /214/ und die ganz tief fallende Kontur von /51/ auf jeweils zwei Silben verteilt. Demnach hat

die erste Silbe einen Wert von /21/ für T3 und /53/ für T4. Die Konturen nach T1 und T2 kann sie mit ihrer Theorie allerdings nur schwer erklären. Daraufhin schlägt sie vor, dass der Neutralton generell das Registermerkmal [-Upper] besitzt. Im Falle einer steigenden Kontur des Neutraltons nach T3 befindet sich die Realisierung des Tonmerkmals [+high] innerhalb des Registermerkmals [-Upper] (siehe Kapitel 2, Abschnitt 2.1.1.1). Auch Peng et al. (2005) beschreiben die Oberflächenkontur des Neutraltons als ein Resultat tonaler Steuerung des davorstehenden vollen Tons. Shen (1992) geht noch einen Schritt weiter und behauptet, dass die letzte tonale Komponente in der vorausgehenden Silbe das Tontarget des Neutraltons vorgibt.

Chen und Xu (2006) untersuchen die phonetische Realisierung des Neutraltons gründlich und stellen die These auf, dass der Neutralton ein grundlegendes Tontarget besitzt, welches sich in der Mitte des Stimmumfangs befindet. In ihrer Studie manipulieren sie die Anzahl der Silben mit Neutralton zwischen zwei Silben mit vollen Tönen.

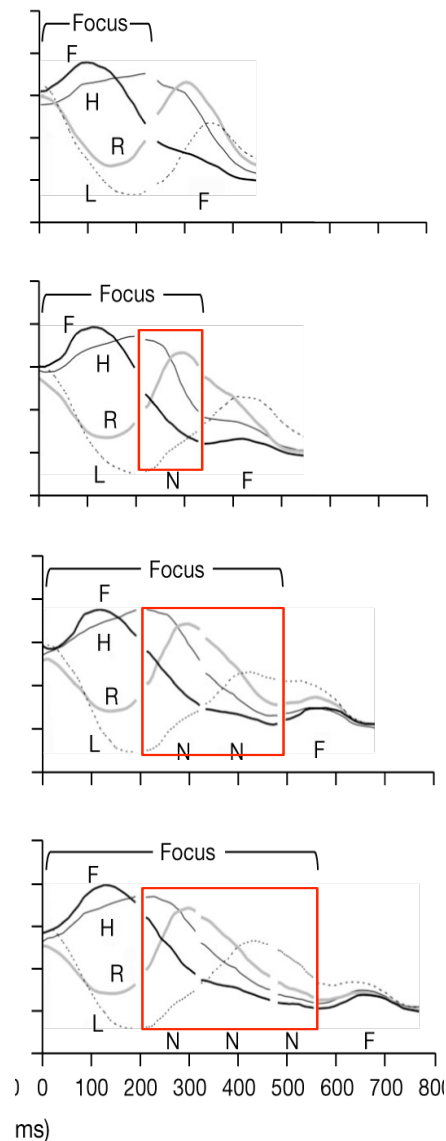


Abbildung 6.1: F0-Konturen von Äußerungen, in denen die erste Silbe einen der vier vollen Töne (H=T1, R=T2, L=T3 und F=T4) und die letzte Silbe den T4 (F) trägt. Die Anzahl der eingeschobenen Silben mit Neutralton (N) zwischen den zwei Silben mit vollen Tönen variiert (Chen & Xu 2006).

Abbildung 6.1 zeigt, dass sich der Einfluss des vorausgehenden Tons auf den Neutralton reduziert, wenn sich die Anzahl der dazwischenstehenden Silben mit Neutralton vergrößert. Am Ende der dritten Silbe mit Neutralton in der letzten Reihe laufen die F0-Konturen des Neutraltons im mittleren Bereich zusammen. Dies deutet darauf hin, dass sich die F0-Kontur des Neutraltons im mittleren Tonbereich einpendelt, wenn ausreichend Zeit vorhanden ist. Dies erfolgt unabhängig davon, welcher volle Ton davor steht. Chen und Xu verwenden das um zu argumentieren, dass der Neutralton im SC tatsächlich ein grundlegendes Target hat. Allerdings unterscheidet sich die Implementierung des Tontargets

stark von den vollen Tönen. Aufgrund der kurzen Dauer und des metrisch schwachen Status wird das phonetische Target des Neutraltons unter dem starken Einfluss des vorausgehenden vollen Tons schwergängig realisiert. Deshalb erreichen die vollen Töne in der Regel ihre Tontargets gegen Ende der Silbe, während der Neutralton sein Target am Ende der Silbe häufig nicht vollständig erreicht (Chen und Xu 2006). Sie finden zusätzlich heraus, dass die hoch variable F0-Kontur eines Neutraltons hauptsächlich durch den Offset-Wert und die finale Geschwindigkeit der F0-Bewegung des vorausgehenden Tons verursacht wird. Gegen letztere hat der Neutralton einen wirkungslosen Widerstand (*inefficient resistance* nach Chen & Xu 2006), d.h. die Kontur eines Neutraltons nach T4 fällt weiter nach unten, obwohl sein eigentliches Target in der Mitte liegt und die Kontur nach dem niedrigen Offset-Wert des T4 wieder steigen sollte.

Zusammenfassend lässt sich sagen, dass der Neutralton sein eigenes phonetisches Tontarget besitzt, sich in der Mitte des Stimmumfangs befindet und statisch zu sein scheint. Die Realisierung des Tontargets wird allerdings vom vorausgehenden Ton stark beeinflusst.

Eine fallende Kontur (z.B. H*L-) kann im Deutschen auf zwei unterschiedliche Arten realisiert werden, ohne ihre phonologische Spezifikation zu verändern. Der Gipfel in der Realisierung von H* kann früh oder spät in der betonten Silbe *männ-*aligniert werden. Die kontrastiven F0-Konturen der Sequenzen T1-T0 und T4-T0 haben ähnliche Oberflächenkonturen wie die zwei Realisierungen derselben tonalen Spezifikation H*L- im Deutschen, sind allerdings bedeutungsunterscheidend. Obwohl sich zweisilbige Wörter mit Neutralton im SC eine sehr ähnliche metrische Struktur mit einem Trochäus im Deutschen teilen, unterscheiden sie sich doch stark bei den zugrundeliegenden Mechanismen der Realisierung der Tontargets. Während Muttersprachler bei der Realisierung der metrisch schwachen Silbe mit dem Neutralton versuchen, das konkrete Tontarget zu erreichen, behandeln deutsche Lerner die metrisch schwache Silbe innerhalb des Trochäus wie eine Interpolation zwischen zwei Tontargets. Aufgrund fehlender Strategien und Trainings bei der Tonproduktion von Wörtern mit

Neutralton im SC ist zu erwarten, dass die Lerner bei einer ähnlichen Struktur im SC die Mechanismen der tonalen Realisierung aus dem Deutschen entlehnen. Demzufolge kann die tonale Kontrastivität unter den Tonsequenzen T1-T0, T2-T0, T3-T0 und T4-T0 gewissermaßen verloren gehen.

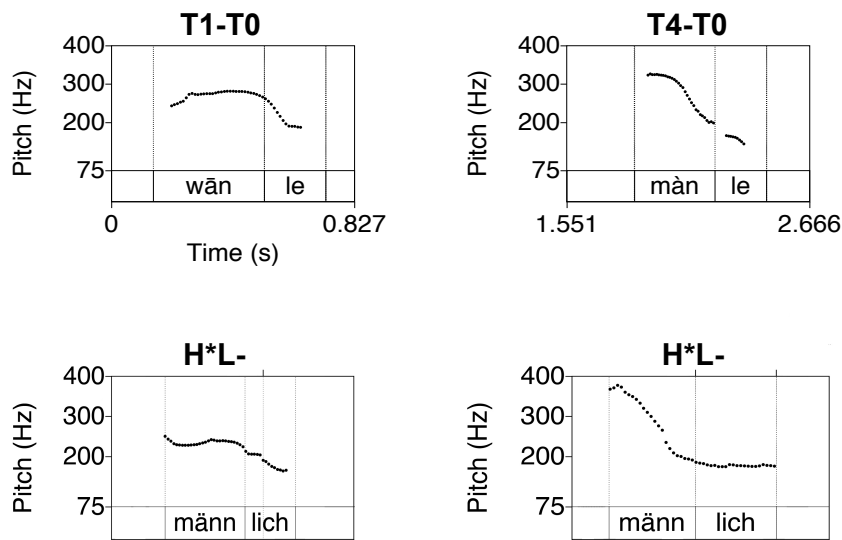


Abbildung 6.2: Die F0-Kontur der Tonsequenzen T1-T0 (*wān le* – „gebogen“) und T4-T0 (*màn le* – „langsam geworden“) sowie die Realisierung von H*+L- im Deutschen.

Des Weiteren wird in der Fachliteratur mehrmals berichtet, dass die tonale Realisierung bestimmter Tonpaare miteinander verschmilzt. Chen (2000:182) berichtet:

Confusion typically occurs between Tone 2 (rising) and Tone 3 (low-level), and between Tone 4 (falling) and Tone 1 (high-level).

[Verwechslungen treten typischerweise bei T2 (steigend) und T3 (tief-eben), sowie bei T4 (fallend) und T1 (hoch-eben) auf.]

(Chen 2000:182, meine Übersetzung)

Sun (1998: 199-200) berichtet ebenfalls von einer Verschmelzung bei der Produktion von T2 und T3. Hao (2012: 273) untersucht die Produktion und

Perzeption der Töne im SC bei Sprechern, deren L1 eine Ton- bzw. Nichttonsprache ist und stellt fest, dass bei Sprechern beider Gruppen die Fehler am häufigsten durch die Verschmelzung/Verwechslung von T2 und T3 sowie T4 und T1 auftreten.

Eine Verwechslung von T2 und T3 berichten auch Li und Thompson (1977) in der Untersuchung muttersprachlicher Kinder beim Tonerwerb:

...the rising and dipping tones are substituted for each other throughout the tone acquisition process.

[...die steigenden und tief-fallenden Töne werden während des Erwerbsprozesses durchgehend gegeneinander ausgetauscht.]

(Li & Thompson 1977: 185, meine Übersetzung)

Das zeigt, dass die Verschmelzung/Verwechslung von bestimmten Tonpaaren nicht nur bei L2-Lernern auftritt, sondern ein generelles Phänomen des Tonerwerbs darstellt. Sun (1998: 199) bezeichnet diese häufig zu beobachtende Schwierigkeit als „*inherent difficulties to the learner*“ (inhärente Schwierigkeiten für Lerner). Demzufolge wird die folgende Hypothese aufgestellt:

H4: Lerner produzieren den Neutralton im SC wie eine unbetonte Silbe im Deutschen, wodurch lexikalische Kontraste in der Tonproduktion zum Teil verloren gehen und die tonale Realisierung bestimmter Tonpaare miteinander verschmilzt.

6.2 Methodik

6.2.1 Testmaterialien

In der vorliegenden Studie wird die Realisierung von vier zweisilbigen Wörtern mit Neutralton untersucht. Die erste Silbe trägt einen der vier vollen Töne, die zweite Silbe den Neutralton und bleibt “子”, ein Nominalsuffix mit intrinsischem Neutralton wie in (6.1):

(6.1)	bēi zi	hóu zi	lǒu zi	mào zi
	杯子	猴子	篓子	帽子
	(Tasse)	(Affe)	(Korb)	(Hut)
	T1-T0	T2-T0	T3-T0	T4-T0

Diese vier sinntragenden Wörter werden jeweils in der Fokus- und Präfokus-Position in Trägersätze mit vier Kontextwörtern²¹ eingebaut, wie in (6.2):

(6.2)	Zielwörter			
	<div style="display: flex; align-items: center; justify-content: center;"> <div style="margin-right: 10px;"> wǒ kàn jiàn 我 看 见 ich sehen-sehen </div> <div style="margin-right: 10px;"> <div style="font-size: 4em; line-height: 1;">{</div> <div style="text-align: center;"> bēi zi 杯 子 hóuzi 猴 子 lǒu zi 篓 子 màozì 帽 子 </div> </div> <div> zài xiǎo wū shàng. 在 小 屋 上。 sein klein Haus auf </div> </div>			

„Ich sehe, dass ein(e) **Zielwort** auf dem kleinen Haus ist.“

²¹ An dieser Position des Trägersatzes werden außer „xiǎowū 小屋“ (ein kleines Haus) noch „xiǎohú 小湖“ (ein kleiner See), „xiǎolǔ 小橹“ (ein kleines Paddel) und „xiǎolù 小路“ (ein kleiner Weg) verwendet. Diese Wörter sind nicht Forschungsgegenstand der vorliegenden Arbeit und werden als „Kontextwörter“ bezeichnet.

Die vier Zielwörter werden in Zitierform und im Trägersatz realisiert. Zusammen ergeben sich 424 Datenpunkte aus der folgenden Berechnung:

[4 Zielwörter (in Zitierform) × 5 Sprecher²² × 2 Gruppen (Lerner vs. Muttersprachler)] + [4 Zierwörtern × 4 Trägersätze mit unterschiedlichen Kontextwörtern × 2 Fokusstatus (in Fokus vs. Präfokus) × 6 Sprecher × 2 Gruppen (Lerner vs. Muttersprachler)].

6.2.2 Ablauf der Datenerhebung

Alle Probanden sitzen bei der Aufnahme in der Sprecherkabine der Abteilung Phonetik des Linguistischen Instituts der Universität zu Köln. Den Probanden wird zuerst vor der Aufnahme Zeit gegeben, sich mit den Zielwörtern vertraut zu machen. Anschließend werden sie aufgefordert, alle Zielwörter laut vorzulesen, die mit einem Bildchen, chinesischen Zeichen und Pinyin angezeigt werden (siehe Abbildung 6.3). Nach einer kurzen Pause werden sie in die zweite Aufgabe eingeführt.



Abbildung 6.3: Vier Zielwörter.

²² Die Realisierungen der Zielwörter in Zitierform von Lerner 3 sind nach der Datenverarbeitung aufgrund eines technischen Problems nicht mehr nutzbar.

Um Tonsequenzen innerhalb bestimmter syntaktischer Strukturen hervorzurufen, wird eine semispontane Leseaufgabe eingesetzt. Diese Aufgabe ist eine Art Lückentest, d.h. den Probanden werden nacheinander Trägersätze präsentiert. Diese Trägersätze enthalten jeweils eine Lücke, in welche die Probanden das passende Zielwort entsprechend dem dazu gezeigten Bild einfügen. Ein Beispiel einer solchen Bildsequenz wird in Abbildung 6.4 illustriert:

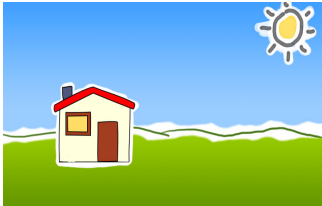
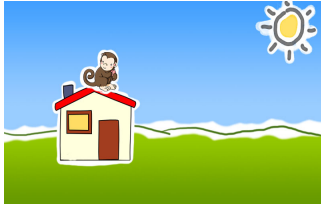
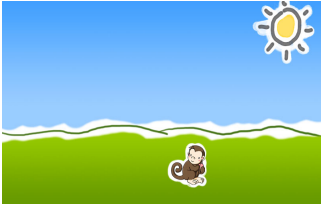
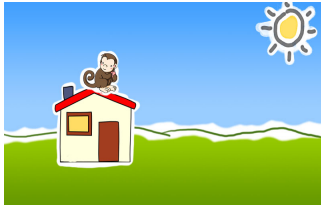
	
<p>Huà lǐ yǒu ____.</p> <p>画 里 有 ____.</p> <p>Bild in existieren</p>	<p>Nǐ kàn jiàn shénme zài xiǎo wū shàng?</p> <p>你 看 见 什 么 在 小 屋 上?</p> <p>du sehen-sehen was sein klein Haus oben</p> <p>‘Was siehst Du auf dem kleinen Haus?’</p>
<p>„Es gibt ____ auf dem Bild.“</p>	<p>wǒ kàn jiàn ____ zài xiǎo wū shàng.</p> <p>我 看 见 ____ 在 小 屋 上.</p> <p>ich sehen-sehen sein klein Haus auf</p> <p>‘Ich sehe ____ auf dem kleinen Haus.’</p>
	
<p>Huà lǐ yǒu ____.</p> <p>画 里 有 ____.</p> <p>Bild in existieren</p>	<p>Nǐ kàn jiàn hóuzi zài shénme shàng?</p> <p>你 看 见 猴 子 在 什 么 上?</p> <p>du sehen-sehen Affe sein was auf</p> <p>‘Worauf siehst Du den Affe sich befinden?’</p>
<p>‘Es gibt ____ auf dem Bild.’</p>	<p>wǒ kàn jiàn hóuzi zài ____ shàng.</p> <p>我 看 见 猴 子 在 ____ 上.</p> <p>ich sehen-sehen Affe sein ____ auf</p> <p>‘Ich sehe den Affe auf ____.’</p>

Abbildung 6.4: Beispiel einer Bildsequenz mit dem Zielwort „hóuzi“ (,der Affe‘). Die Interlineare Morphem - Glossierung und die deutsche Übersetzungen sind nachträglich hinzugefügt.

Insgesamt werden 16 solcher Bildsequenzen jeweils mit zwei Bildpaaren eingesetzt. Das linke Bild eines Paares enthält einen einführenden Satz und das rechte einen Satz mit dem Zielwort. Das Zielwort wird im oben rechten Bild (in Abbildung 6.4) im Fokus und im unten rechten Bild im Präfokus realisiert.

Alle 16 Bildsequenzen werden in einer PowerPoint-Präsentation und in zufälliger Reihenfolge gezeigt. Wenn ein Proband Fehler im segmentalen Bereich macht

oder einen Satz mit auffällig unpassenden Pausen produziert, wird er aufgefordert, diesen Satz zu wiederholen.

Die F0-Extraktion und Normalisierung wird mit dem gleichen Verfahren durchgeführt, wie in Kapitel V Abschnitt 5.2.3 beschrieben.

6.3 Ergebnisse

6.3.1 Zielwörter realisiert in Zitierform

F0-Konturen

Zuerst wird die Tonproduktion der Zielwörter in Zitierform von Lernern und Muttersprachlern verglichen. In Abbildung 6.5 werden alle Realisierungen von Zielwörtern in Zitierform dargestellt. Jede produzierte Silbe wird in Praat (Boersma & Weenink 2013) analysiert. Mit Hilfe eines Praat-Skriptes (Xu 2013) werden die F0-Werte vom stimmhaften Anteil einer annotierten Silbe gemessen und die Werte an zehn gleichabständigen Punkten extrahiert. In der linken Spalte ist die Produktion von Lernern zu sehen und in der rechten Spalte die der Muttersprachler. Jede Farbe stellt einen Sprecher dar. Während bei den Lernern sowohl in der Tonhöhe als auch den Konturen der Töne viel Variation zu sehen ist (bis auf die Sequenz T4-T0), zeigen die Muttersprachler bei den Konturen der Töne ein einheitliches Muster. Bei der Beobachtung der Realisierung des Neutraltons ist besonders auffällig, dass der Lerner DL15 die Silbe des Neutraltons mit einer eindeutig fallend-steigenden Kontur in den Tonsequenzen T1-T0, T2-T0 und T3-T0 produziert. Auch Lerner DL11 zeigt dieses Muster bei den Realisierungen von T1-T0 und T2-T0. Bezüglich der Kontur der vollen Töne ist bei T2 die größte Variation zu erkennen. Insgesamt zeigen die Lerner drei unterschiedliche Konturen bei der Realisierung des T2: hoch-eben, steigend und fallend-steigend, während alle Muttersprachler eine leicht fallende und dann wieder steigende Kontur zeigen. Auch die Realisierungen von T3 in der Sequenz T3-T0 unterscheiden sich stark bei Lernern und Muttersprachlern. Lerner

produzieren diesen Ton mit einer fallend-steigenden Kontur, was der Grundform des T3 entspricht, während die Muttersprachler diesen Ton mit der so genannten halben T3-Realisierung produzieren, d.h. T3 nur mit einer fallenden Kontur realisieren, wenn er an einer nicht finalen Position steht.

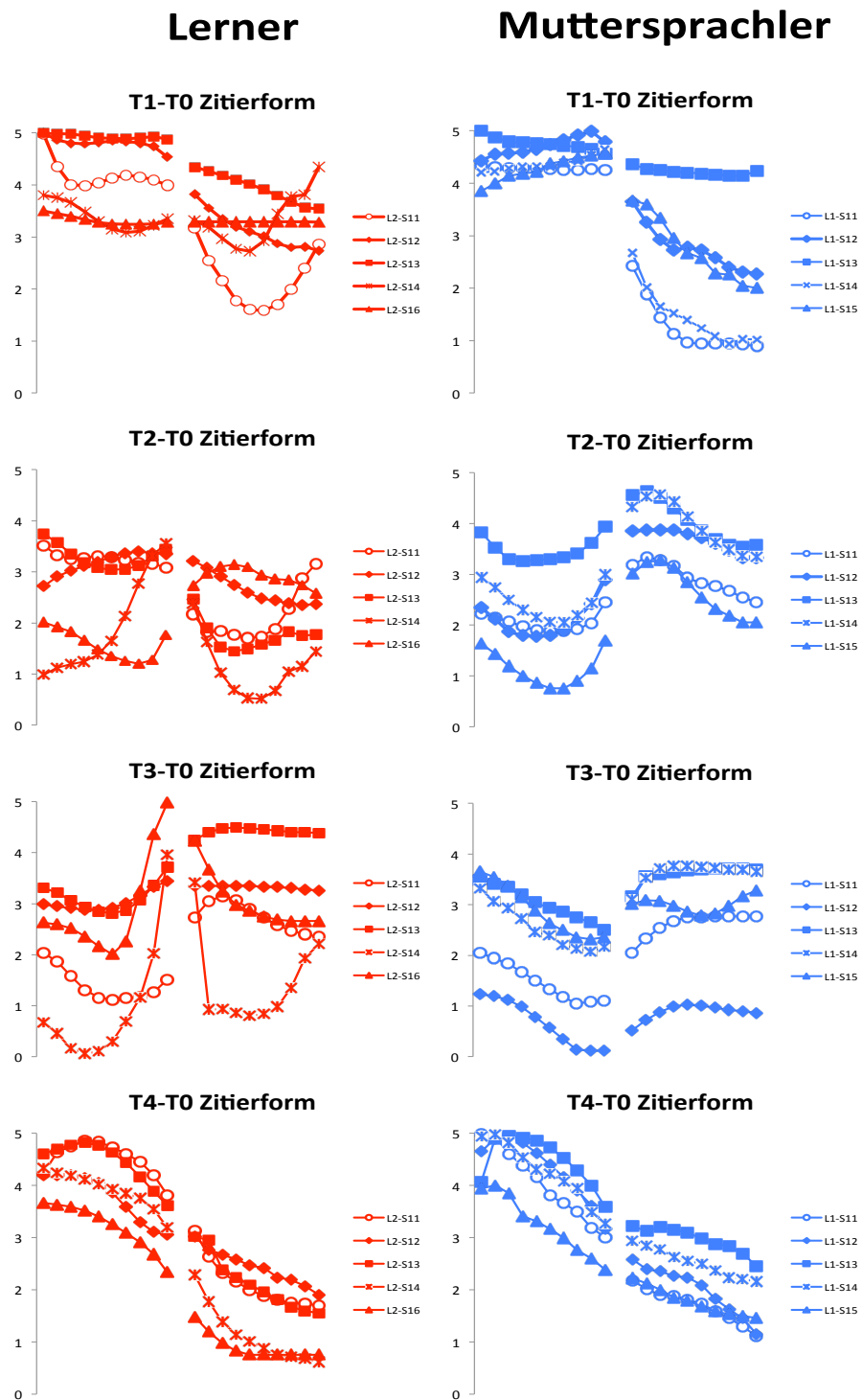


Abbildung 6.5: Individuelle Darstellung der Realisierung von Wörtern mit Neutralton in Zitierform durch Lerner (linke Spalte) und Muttersprachler (rechte Spalte).

Der F0-Wert jedes einzelnen Messpunkts wird mittels eines *mixed linear regression model* vorhergesagt. Dabei ist der F0-Wert des jeweiligen Punkts die abhängige Variable, die GROUP (Lerner vs. Muttersprachler) der *fixed* Faktor und das *intercept* von jedem Sprecher (Speaker) ist der *random* Faktor.

(6.3)

$$\text{F0_Wert} \sim \text{GROUP} + (1 \mid \text{Speaker})$$

(von jedem Messpunkt)

Demzufolge erhält man 20 *mixed linear regression models* und den jeweiligen Wert der oberen 95%- und der unteren 5%-Vertrauensgrenze (*confidence interval*) für jeden Punkt. Durch die obere und untere Vertrauensgrenze lässt sich eine Zone festlegen, in der alle Werte der Messpunkte abhängig von der Identität des vollen Tons und der Gruppe höchstwahrscheinlich liegen (siehe Abbildung 6.6). Somit lassen sich bestimmte Muster voraussagen, die ein Lerner oder Muttersprachler bei der Produktion einer bestimmten Tonsequenz am wahrscheinlichsten zeigt.

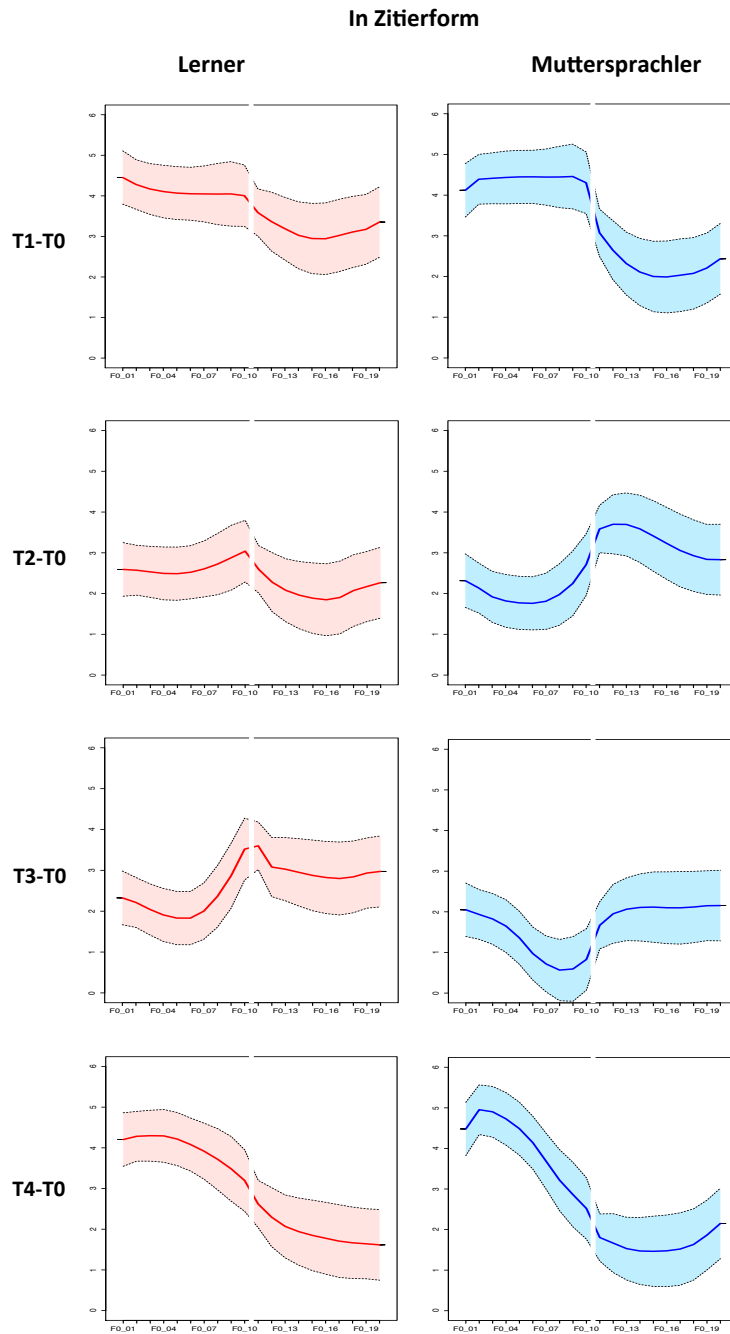


Abbildung 6.6: Die obere und untere Vertrauensgrenze der Realisierung verschiedener Tonsequenzen in Zitierform bei verschiedenen Sprechergruppen (rot = Lerner; blau = Muttersprachler). Die durchgezogenen Linien stellen den Mittelwert dar.

In Abbildung 6.6 zeigen Lerner und Muttersprachler ähnliche Muster bei der Realisierung der Tonsequenz T1-T0. Sprecher beider Gruppen realisieren die erste Silbe dieser Tonsequenz mit einer hoch-flachen Kontur, die dann in der zweiten Silbe abfällt. Bei T2-T0 realisieren die Muttersprachler T2 mit einer steigenden Kontur, dessen Anstieg bis zum Anfang der nächsten Silbe mit Neutralton anhält, wo er den Gipfel der Kontur erreicht und schließlich innerhalb der zweiten Silbe

abfällt. Lerner zeigen hingegen eine relativ flache Kontur bei der Silbe mit T2, die erst gegen Ende der Silbe ansteigt. Die Kontur des Neutraltons ist leicht fallend-steigend, was der Kontur des Neutraltons in der Sequenz in T1-T0 ähnlich ist. Muttersprachler realisieren T3-T0 mit einer tief-fallenden Kontur für T3, die dann mit einem „Sprung“ in der Kontur zu der nächsten Silbe mit Neutralton hin ansteigt. Die Realisierung dieser Tonsequenz durch Lerner zeigt zuerst eine fallend-steigende Kontur für T3, die der kanonischen Form des T3 entspricht, während die Kontur der Silbe mit Neutralton die Tonhöhe des Offsets von dem vorausgehenden Ton behält und flach bleibt. Die Tonsequenz T4-T0 wird von Lernern und Muttersprachlern mit einer sehr ähnlichen Kontur realisiert: Eine fallende Kontur bei der Silbe mit T4 und eine tiefe, relativ flache Kontur bei der zweiten Silbe.

Zusammenfassend ist festzustellen, dass der entscheidende Unterschied zwischen Lernern und Muttersprachlern bei der tonalen Realisierung im Bereich des Übergangs zwischen den zwei Silben liegt. Am deutlichsten ist dieser Unterschied am Onset der Silbe mit Neutralton. Die Ergebnisse der oben beschriebenen 20 *mixed linear regression models* werden in Tabelle 6.2 zusammengefasst.

Messpunkte	Signifikanz (<i>p-value</i>) des Unterschieds zwischen DL und CM
F0_01	0.5512
F0_02	0.06788
F0_03	0.05052
F0_04	0.06096
F0_05	0.03862 *
F0_06	0.01512 *
F0_07	0.01206 *
F0_08	0.01158 *
F0_09	0.006289 **

F0_10	0.000713 ***
F0_11	2.372e-05 ***
F0_12	0.0009401 ***
F0_13	0.001098 **
F0_14	0.002688 **
F0_15	0.007453 **
F0_16	0.01737 *
F0_17	0.04835 *
F0_18	0.1071
F0_19	0.1488
F0_20	0.1567

Tabelle 6.2: Ergebnisse der *mixed linear regression models* beim Vergleich der F0-Werte der Messpunkte zwischen Lernern und Muttersprachlern.

In Tabelle 6.2 lässt sich feststellen, dass es signifikante Unterschiede zwischen Lernern und Muttersprachlern bei der Realisierung aller vier Tonsequenzen gibt, die konzentriert bei dem Übergang zwischen den Silben auftreten.

In der muttersprachlichen Realisierung beeinflusst der Offset-Wert des vorausgehenden vollen Tons den Onset-Wert des Neutraltons stark, sodass sich die Onset-Werte des Neutraltons in unterschiedlichen Tonsequenzen deutlich voneinander unterscheiden. Beispielsweise zeigt der Neutralton bei T2-T0 eine fallende und bei T3-T0 eine leicht steigende Kontur. Dagegen zeigen Lerner eine ähnliche Kontur des Neutraltons bei der Realisierung in den Sequenzen T2-T0 und T3-T0 und die Offset-Werte der vollen Töne scheinen von den Onset-Werten des nachfolgenden Neutraltons beeinflusst zu sein.

Die Konturen von Tonsequenzen mit einem Neutralton lassen bei der muttersprachlichen Realisierung auf vorwiegend progressive tonale Koartikulation schließen, die ebenfalls in der Produktion von Tonsequenzen mit vollen Tönen zu finden ist. Nach Chens und Xus Theorie hat der Neutralton ein grundlegendes phonetisches Target, dürfte das Target des Neutraltons dem progressiven Effekt der tonalen Koartikulation in SC unterliegen. Demzufolge

können die Oberflächenkonturen des Neutraltons als Resultate besonders ausgeprägter progressiver Koartikulation interpretiert werden. In der Lerner-Produktion ist dieses Muster nicht zu finden.

Es gibt einige mögliche Erklärungen für die oben genannten Phänomene. In erster Linie wird die starke Variation in der Realisierung des Neutraltons nach den vier vollen Tönen im Unterricht nicht explizit behandelt und geübt. Wie in Kapitel 3 beschrieben, werden Silben mit einem Neutralton sowohl im Lehrmaterial als auch im Unterricht als kurze und unbetonte Silben gekennzeichnet. Sie erhalten zudem keinen Ton. Wie in den oben dargestellten Daten zu sehen ist, greifen die Lerner in Ermangelung einer Anweisung zur phonetischen Produktion des Neutraltons auf zwei Strategien zurück. DL11 und DL15 realisieren die Silbe mit dem Neutralton *zi*, der orthographisch als *zi* dargestellt wird, mit der Kontur eines T3. Dies kommt daher, dass dieses Zeichen im SC zwei Lesarten hat. Es hat einen intrinsischen Neutralton, wenn es als ein Nominalsuffix fungiert, so wie es in der vorliegenden Studie der Fall ist. Es könnte jedoch auch „der Sohn“ bedeuten und hätte dann einen T3. Die zwei Lerner haben sich für die letztere Aussprache des Heteronyms entschieden, da sie in der Regel nur in der Produktion isolierter Töne geschult werden. Bei der Zusammensetzung von Tönen versuchen sie nach wie vor, die Grundformen der Töne zu produzieren und damit die nötige Kontrastivität zu gewährleisten. Bei einem Neutralton existiert allerdings keine Grundform, auf die Lerner zurückgreifen können. Deswegen versuchen sie mit anderen Mitteln die Kontrastivität zu erreichen. Ein Heteronym bietet sich als gute Alternative an.

Andere Lerner versuchen, die Silbe mit dem Neutralton wie eine unbetonte Silbe zu behandeln, wie es ihnen in den Lehrbüchern und im Unterricht vermittelt wurde. In Wirklichkeit ist eine Silbe mit einem Neutralton aber nicht mit einer unbetonten Silbe im Deutschen gleichzusetzen. Sie teilen sich die Eigenschaft, dass sie kurz und metrisch schwach sind. Allerdings hat ein Neutralton im SC ein phonetisches Target und eine unbetonte Silbe im Deutschen dient als Interpolation zwischen zwei betonten Silben und besitzt entsprechend kein phonetisches Target.

Bei der Realisierung von T3-T0 steigt die Kontur nach einem L* in der unbetonten Silbe, die einer Kontur L*+H im Deutschen sehr ähnelt. Grice und Kollegen (2012) stellen fest, dass die Kontur L*+H von deutschen Sprechern bevorzugt zur Markierung des Verum-Fokus aber seltener als lexikalischer Kontrast eingesetzt wird. In Abbildung 6.5 ist deutlich zu erkennen, dass die Lerner (außer DL11) die Sequenz T3-T0 nicht mit der L*+H Kontur realisieren.

Dauerverhältnis

Insgesamt zeigen Lerner und Muttersprachler ein ähnliches Dauerverhältnis bei der Tonproduktion. Lediglich bei DL11 kommt es zu starken Abweichungen beim Dauerverhältnis, wie in Abbildung 6.7 dargestellt. Er produziert die Silbe mit dem Neutralton fast gleich lang wie die Silbe mit einem vollen Ton. Allerdings zeigt DL11 in Abbildung 6.5 bei der Realisierung der Sequenz T3-T0 eine Kontur, die der muttersprachlichen ähnlich ist. Da das Dauerverhältnis bei der Realisierung von Silben mit Neutralton essentiell ist, realisiert DL11 die Tonsequenz T3-T0 trotz ähnlicher Kontur nicht muttersprachlich.

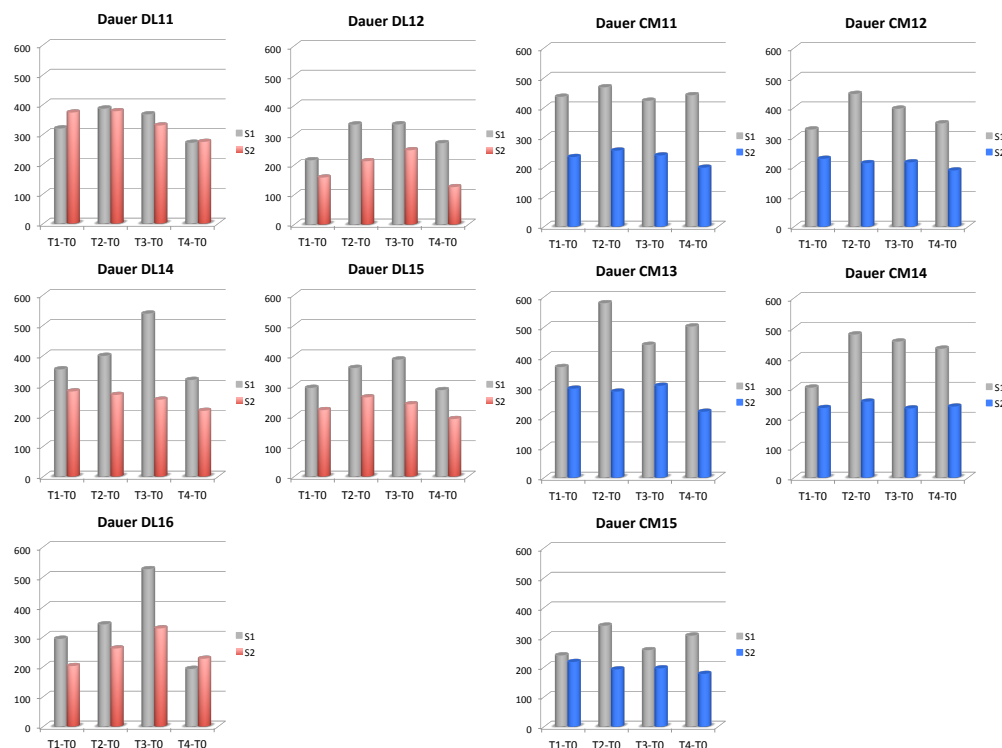


Abbildung 6.7: Dauerverhältnis zwischen der Silbe mit vollem Ton und Neutralton in Zitierform. Die grauen Balken stellen die Dauer der Silben mit vollen Tönen dar (ms). Die bunten Balken zeigen die Dauer der Silben mit Neutralton: Rot für Lerner, Blau für Muttersprachler.

Insgesamt zeigen die Lerner bei der Produktion eines Neutraltons in zweisilbigen Wörtern in Zitiertform gewisse Schwierigkeiten. Sie versuchen jedoch, die lexikalische Kontrastivität zu gewährleisten, indem sie die Töne in ihren Grundformen produzieren. Daraus ergibt sich die Frage, was passiert, wenn diese Wörter in einer größeren Domäne realisiert werden, z.B. in einer Äußerung. Im nächsten Abschnitt werden deshalb die Ergebnisse der tonalen Realisierung eines Neutraltons auf der Äußerungsebene dargestellt.

6.3.2 Zielwörter realisiert in einer Äußerung

F0-Konturen

In Abbildung 6.8 sind alle Realisierungen der Zielwörter in Äußerungen von Lernern und Muttersprachlern aufgezeichnet. Rote Linien stehen für Lerner, blaue für Muttersprachler. Durchgezogene Linien stellen die Realisierung der Zielwörter im Fokus und gestrichelte die Realisierung außerhalb des Fokus dar. In jeder Zeile sind Konturen einer Tonsequenz dargestellt.

Zuerst wird die Realisierung einer der vier Tonsequenzen bei verschiedenen Sprechern beider Gruppen untersucht. In der ersten Zeile kann man sehen, dass alle Muttersprachler T1 mit einer hohen Kontur produzieren. Diese Kontur steigt ganz leicht bis zum Ende der Silbe und erreicht dann den höchsten Punkt. Ab Beginn der zweiten Silbe fällt die Kontur stetig bis zum Silbenende. Außer DL11 zeigen alle Lerner eine abweichende Form. DL12, 13, 14 und 15 realisieren T1 mit einer eindeutig fallenden Kontur, die in der Silbe mit T0 noch weiter fällt. DL16 zeigt als einziger Lerner eine vollständig andere Kontur. Sie ist tief in der ersten und hoch-fallend in der zweiten Silbe.

Die Muttersprachler zeigen bei der Sequenz T2-T0 eine deutlich steigende Kontur in der Silbe mit T2, deren Anstieg bis zum Anfang der nächsten Silbe mit T0 anhält. Dies entspricht der Implementierung des Tontargets von T2 im SC, wie es in der Literatur (z.B. in Xu & Wang 2001) berichtet wird. Bemerkenswert ist, dass das Onset der Silbe mit T0 immer höher als das Offset der Silbe mit T2 ausfällt.

Die Lerner zeigen allerdings bei dieser Sequenz viel Variation. Nur DL15 und DL16 realisieren diese Tonsequenz durchgehend mit einer muttersprachlichen Kontur. DL12 realisiert diese Tonsequenz mit sehr unterschiedlichen Konturen, so wie auch DL14.

Die Sequenz T3-T0 zeigt bei allen Muttersprachlern eine tief-fallende Kontur in der Silbe mit T3. Dann macht die F0-Bewegung einen relativ abrupten Sprung an den Anfang der Silbe mit T0. Die Realisierungen der Lerner erfolgen bei dieser Sequenz ebenfalls sehr variabel. Bis auf DL16 zeigt kein Lerner eine muttersprachliche Kontur. Erneut fallen die Konturen derselben Sequenz bei DL12 sehr unterschiedlich aus, wie auch bei DL14.

Die Sequenz T4-T0 realisieren die Muttersprachler mit einer klar fallenden Kontur in der Silbe mit T4, die in der Silbe mit T0 bis zum Silbenende weiter fällt. Bei den Lernern sind die Realisierungen der Sequenz T4-T0 denen der Muttersprachler am ähnlichsten. Bis auf DL16 zeigen alle Lerner eine fallende Kontur in der Silbe mit T4, die auch in der Silbe mit T0 weiter fällt.

Insgesamt zeigen Muttersprachler bei der Produktion der Zielwörter hinsichtlich der Kontur ein relativ einheitliches Bild, obwohl die Auslenkung bei Anstieg oder Fall der Konturen unter den Muttersprachlern stark variiert. Auffällig ist, dass alle Lerner die Sequenzen T2-T0 und T3-T0 mit ähnlichen oder gleichen Konturen realisieren, obwohl die Muttersprachler diese zwei Sequenzen klar durch die Konturen kontrastieren.

Nun wird die Realisierung der vier Tonsequenzen von jedem einzelnen Sprecher untersucht. Man kann feststellen, dass Muttersprachler die vier Tonsequenzen mit vier kontrastiven Konturen realisieren. Dagegen ist auffällig, dass alle Lerner die Tonsequenzen mit einer reduzierten Zahl kontrastiver Konturen realisieren. Einige Lerner realisieren alle Tonsequenzen mit einer einzigen Kontur, z.B. DL11 und DL16.

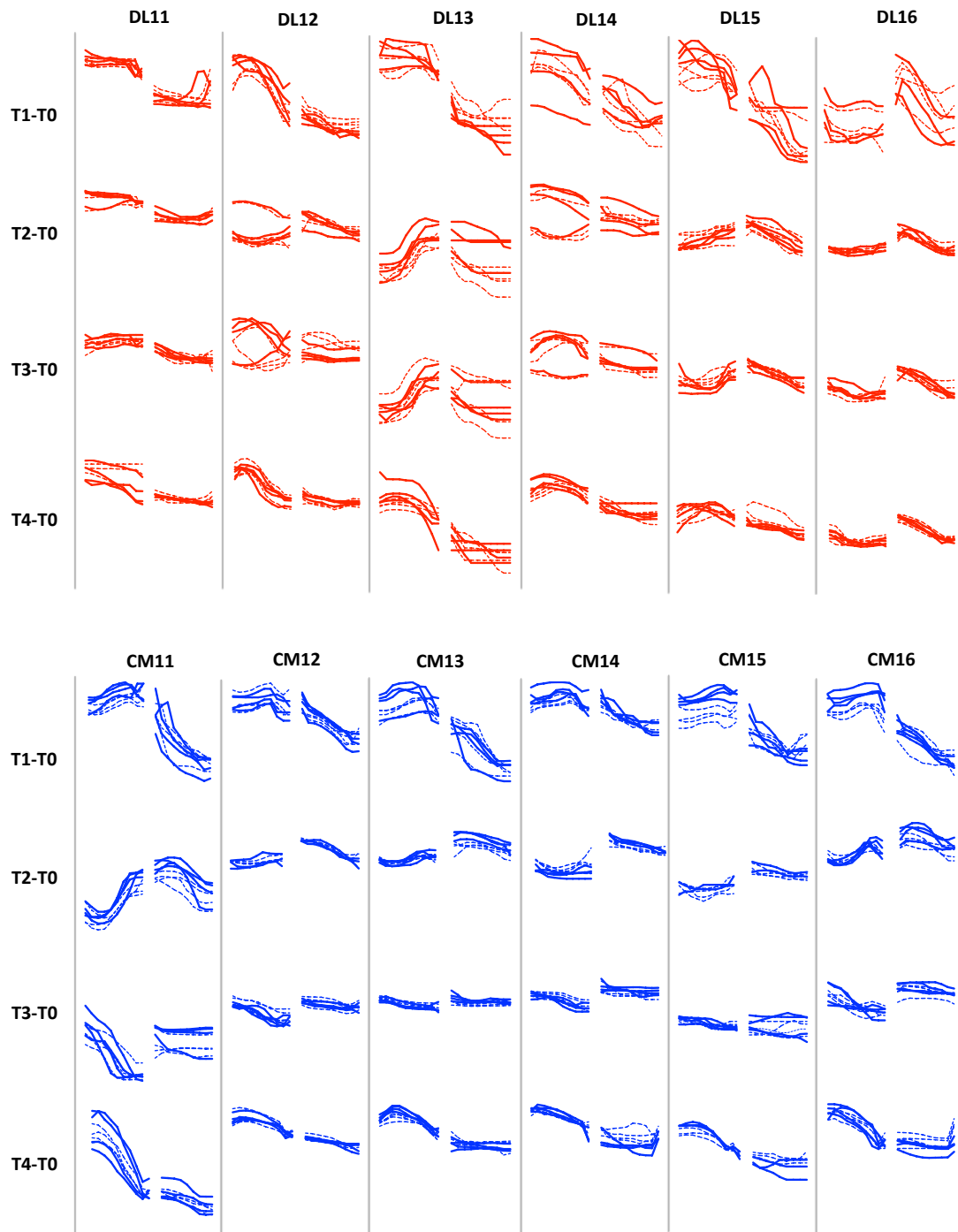


Abbildung 6.8: Konturen zweisilbiger Wörter mit Neutralton von Lernern (rot) und Muttersprachlern (blau) in Äußerungen realisiert.

Um aus der Menge der individuellen Variationen eine Haupttendenz zu generieren, werden die F0-Werte aller 20 Messpunkte mithilfe *mixed linear regression models* vorhergesagt. Die abhängige Variable ist der Wert der *T-value* des einzelnen Messpunkts. Die GROUP (Lerner vs. Muttersprachler) ist der *fixed* Faktor und das *intercept* jedes Sprechers (Speaker) ist der *random* Faktor. Wie in

Abschnitt 6.3.1 beschrieben, erhält man 20 einzelne *mixed linear regression models* und den jeweiligen Wert der oberen 95%- und der unteren 5%-Vertrauensgrenze (*confidence interval*) für jeden Messpunkt. Wenn man die Werte der Vertrauensgrenzen aufzeichnet, erhält man die Haupttendenz, die Lerner und Muttersprachler bei den Realisierungen zeigen (Abbildung 6.9).

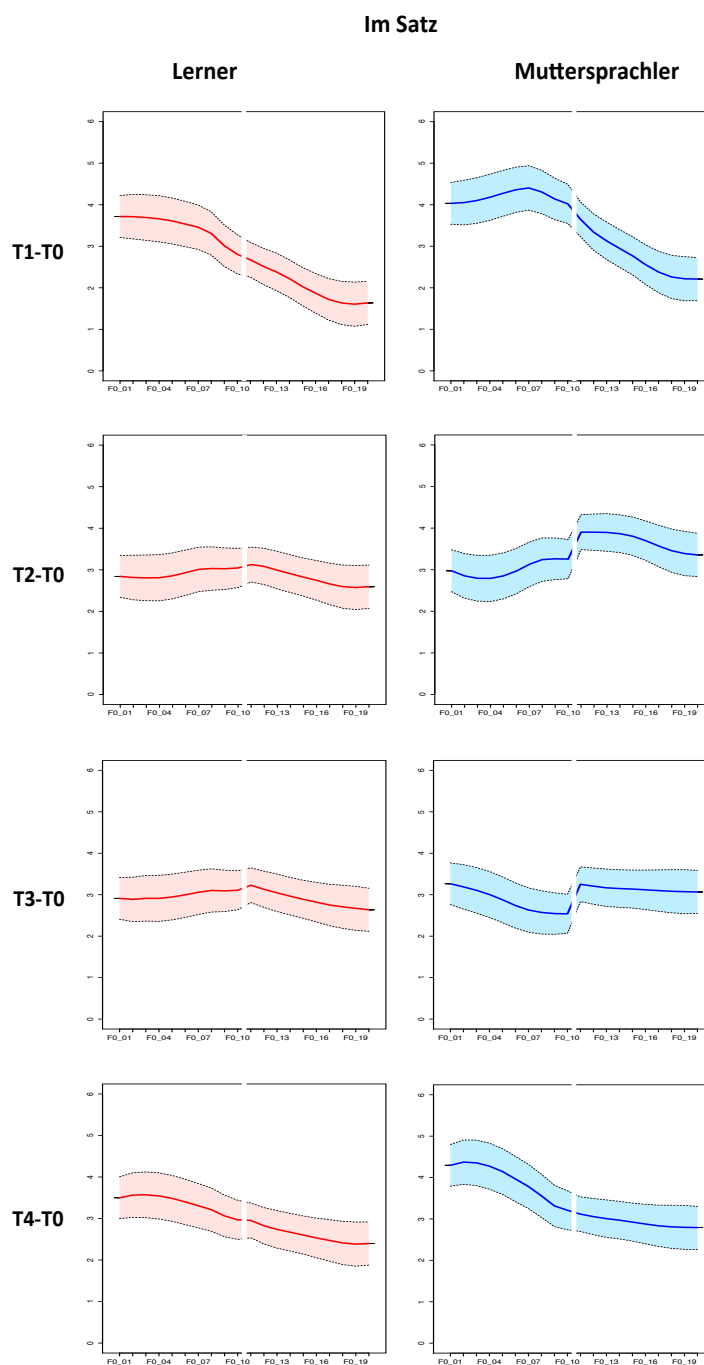


Abbildung 6.9: Vertrauensgrenzen für Realisierungen der in Äußerungen produzierten Zielwörter
Rot steht für Lerner und Blau für Muttersprachler.

Bei der Tonsequenz T1-T0 realisieren Muttersprachler die Silbe mit T1 mit einer hoch-flachen Kontur und die Silbe mit T0 mit einer fallenden, während die Lerner diese Tonsequenz mit einer durchgehenden fallenden Kontur realisieren. Im Vergleich zur Realisierung der Zielwörter in Zitierform realisieren die Lerner die Tonsequenz T1-T0 in Äußerungen mit einer weniger muttersprachlichen Kontur.

Muttersprachler produzieren die Tonsequenz T2-T0 mit einer steigenden Kontur auf der Silbe mit T2, die mit einen „Sprung“ in die Silbe mit Neutralton übergeht. Lerner dagegen realisieren die Silbe mit T2 bei dieser Tonsequenz mit einer flach-steigenden und die Silbe mit Neutralton mit einer flach-fallenden Kontur.

Bei T3-T0 zeigen Muttersprachler eine eindeutig fallende Kontur in der Silbe mit T3 und diese steigt abrupt in die nächste Silbe. Lerner zeigen allerdings eine sehr ähnliche Kontur wie die bei der Realisierung der Tonsequenz T2-T0.

Tonsequenz T4-T0 ist die einzige Sequenz, die Lerner sowohl in der Zitierform als auch in Äußerungen mit einer muttersprachlichen Kontur realisieren.

Wie schon bei der Realisierung der Zielwörter in Zitierform, wird der Onset-Wert des Neutraltons in der muttersprachlichen Produktion in Äußerungen ebenfalls stark von dem Offset-Wert des vorausgehenden Tons beeinflusst, was progressiver tonaler Koartikulation entspricht. Der Hauptunterschied zwischen den Konturen von Lernern und Muttersprachlern liegt auch bei der Produktion von Zielsilben in Äußerungen im Bereich des Übergangs von dem vollen Ton zum Neutralton. Die Ergebnisse des oben beschriebenen *mixed linear regression models* aller 20 Messpunkte der einzelnen Tonsequenz werden in Tabelle 6.3 zusammengefasst.

Messpunkte	Signifikanz (<i>p-value</i>) des Unterschieds zwischen DL und CM			
	T1-T0	T2-T0	T3-T0	T4-T0
F0_01	0.2432	0.7922	0.4342	0.05607
F0_02	0.3125	0.9361	0.5254	0.05939
F0_03	0.2402	0.9922	0.6882	0.07712
F0_04	0.1415	0.9834	0.8507	0.1041
F0_05	0.06532	0.9996	0.896	0.1371
F0_06	0.02357 *	0.9272	0.6139	0.2104
F0_07	0.009337 **	0.772	0.04053*	0.3043
F0_08	0.006652 **	0.5824	0.02942*	0.4651
F0_09	0.002582 **	0.4889	0.02729*	0.6001
F0_10	0.0006906 ***	0.4928	0.02484*	0.5956
F0_11	0.001415 **	0.02614 *	0.009503**	0.7027
F0_12	0.008381 **	0.02227 *	0.0176*	0.6032
F0_13	0.0113 *	0.01569 *	0.01845*	0.5627
F0_14	0.01275 *	0.01236 *	0.019423*	0.5245
F0_15	0.01159 *	0.01194 *	0.5368	0.503
F0_16	0.02215 *	0.01759 *	0.4599	0.4854
F0_17	0.0379 *	0.02809 *	0.395	0.4805
F0_18	0.06394	0.04243 *	0.3665	0.4507
F0_19	0.08524	0.05842	0.3386	0.437
F0_20	0.1304	0.08009	0.2997	0.08704

Tabelle 6.3: Ergebnisse der *mixed linear regression models* beim Vergleich der Werte der Messpunkte zwischen der Realisierung der Zielwörter in Äußerungen bei Lernern und Muttersprachlern.

Anhand der Vertrauensgrenze (Abbildung 6.9) wird deutlich, dass die Muttersprachler auch für die vier Tonsequenzen distinktive Konturen einsetzen. Dagegen zeigen die Lerner hauptsächlich zwei Konturen. Der Kontrast zwischen den Sequenzen T1-T0 und T4-T0 sowie der zwischen T2-T0 und T3-T0 geht verloren. In den dargestellten Ergebnissen in Abbildung 6.9 ist zudem auffällig, dass Lerner die Tonsequenzen, deren Onsets hoch sind, mit einer durchgehend fallenden Kontur über zwei Silben produzieren (in T1-T0 und T4-T0). Die

Tonsequenzen mit einem tiefen Onset (in T2-T0 und T3-T0) werden mit einer steigenden Kontur in der ersten und einer fallenden in der zweiten Silbe produziert.

Dauerverhältnis

Die Dauer, oder besser gesagt das Dauerverhältnis zwischen der Silbe mit einem vollen Ton und der mit einem Neutralton ist ein essentielles Merkmal des Neutraltons.

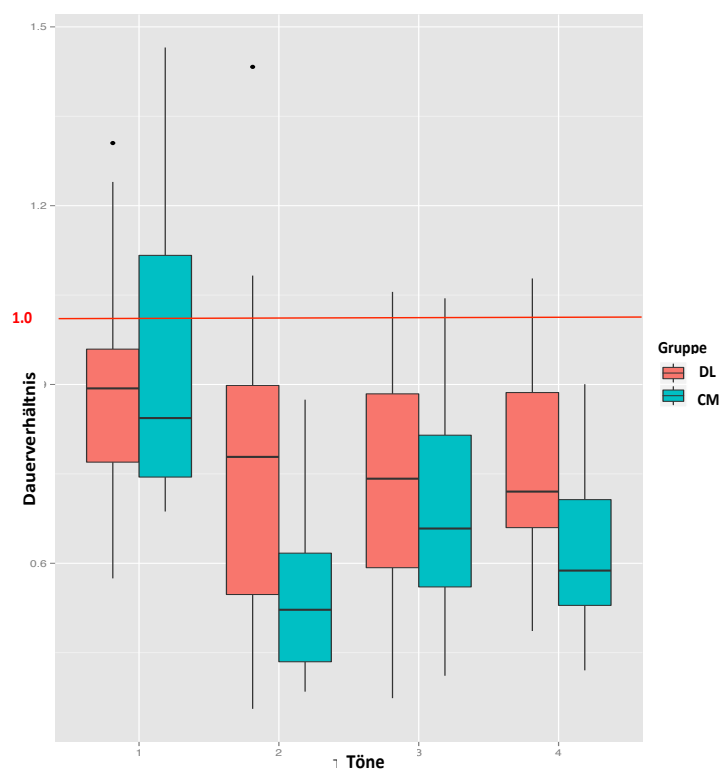


Abbildung 6.10: Das Dauerverhältnis zwischen den zwei Silben in Wörtern mit Neutralton hinter T1, T2, T3 und T4 (1,2,3,4). "DL" steht für Lerner und "CM" steht für Muttersprachler.

Wie in Abbildung 6.10 dargestellt, wird das Dauerverhältnis der zwei Silben in der Produktion der Zielwörter in Äußerungen berechnet und verglichen. Auf der X-Achse repräsentieren die Ziffern 1 bis 4 die vollen lexikalischen Töne T1, T2, T3 und T4, die in einem zweisilbigen Wort dem Neutralton vorausgehen. Die Ziffern auf der Y-Achse repräsentieren das Dauerverhältnis der betreffenden Silben. Man kann sehen, dass, sich das Dauerverhältnis der zwei Silben bei Lernern (rote Kästchen) und Muttersprachlern (blaue Kästchen) geringfügig unterscheidet und

ungefähr bei 1.0 liegt, wenn ein Neutralton nach einem T1 produziert wird. Wird der Neutralton aber nach T2, T3 oder T4 realisiert, gibt es einen klaren Unterschied zwischen den zwei Gruppen bei dem Dauerverhältnis - vor allem nach T2 und T4. Bei den Muttersprachlern ist die Silbe mit dem Neutralton am kürzesten, wenn sie nach T2 produziert wird. Im Vergleich fällt bei den Lernern das Dauerverhältnis bei T2-T0, T3-T0 und T4-T0 nahezu gleich aus.

6.4 Zusammenfassung der Ergebnisse

In diesem Kapitel wurden tonale Realisierungen zweisilbiger Wörter mit einem Neutralton in Zitierform und in Äußerungen untersucht. Sowohl in Zitierform als auch in Äußerungen zeigen die Muttersprachler bei vier unterschiedlichen Tonsequenzen vier kontrastive Konturen. Die Lerner hingegen zeigen unter diesen beiden Bedingungen geringere Kontrastivität als die Muttersprachler. In der Zitierform verschmelzen die Konturen von T2-T0 mit T3-T0 bei der Realisierung durch die Lerner. In Äußerungen verschmelzen zusätzlich noch die Konturen von T1-T0 mit T4-T0. Lerner und Muttersprachler zeigen bezüglich des Dauerverhältnisses das einheitliche Muster, dass die Dauer der Realisierung von Silben mit einem Neutralton bei beiden Gruppen eindeutig kürzer ist, als mit einem vollen Ton.

7 Zusammenfassung, Diskussion und Ausblick

In diesem Kapitel werden die Ergebnisse der beiden empirischen Studien zusammengefasst. Anhand dieser Ergebnisse werden die Hypothesen geprüft und diskutiert. Daraus wird ein Modell des Tonerwerbs deutscher ChaF-Lerner entwickelt. Anschließend werden die didaktischen Konsequenzen der Befunde und zukünftige Forschungsperspektiven erläutert.

7.1 Zusammenfassung der Ergebnisse

Die Ergebnisse der beiden Studien der vorliegenden Arbeit können den zwei wissenschaftlichen Bereichen Tonerwerb und tonale Koartikulation zugeordnet werden.

7.1.1 Ergebnisse im Bereich Tonerwerb

Die erste empirische Studie untersucht die Tonproduktion deutscher ChaF-Lerner bei zweisilbigen Wörtern mit zwei vollen Tönen und zeigt, dass ihre Produktion der Töne in Zitierform nur geringfügig von der der untersuchten Muttersprachler abweicht. Bei Tönen in Sequenzen setzen Lerner jedoch von Muttersprachlern grundlegend abweichende Mechanismen ein. In dem Übergangsbereich zwischen zwei Tönen zeigen die deutschen Lerner vorwiegend regressive tonale Koartikulation, während Muttersprachler erwartungsgemäß vorwiegend progressive tonale Koartikulation zeigen. Zusätzlich ist bei deutschen ChaF-Lernern bei tonaler Koartikulation eine stark dissimilatorische Tendenz festzustellen. Diese Tendenz zeigt sich zum einen dadurch, dass der Offset-Wert eines betreffenden Tons ansteigt, wenn er vor einem Ton mit niedrigem Onset-Wert produziert wird. Zum anderen ist die dissimilatorische Tendenz in der außerordentlich langen Pause zwischen dem ersten und zweiten Ton zu erkennen, die Lerner bei der Produktion von Tonsequenzen in Äußerungen einlegen.

Das Ergebnis der Studie in Kapitel 5 zeigt zudem, dass im ChaF-Unterricht explizit geübte Tonsequenzen von den Lernern meistens zielsprachengerecht produziert werden.

Die zweite Studie untersucht die Tonproduktion deutscher Lerner bei zweisilbigen Wörtern mit einem Neutralton im SC. Sie deckt auf, dass tonale Kontraste bei der Tonproduktion der Lerner zum Teil verloren gehen und die produzierten Konturen bei bestimmten Tonpaaren verschmelzen: Wenn Zielwörter in Zitiertform realisiert werden, sind die Konturen der Tonpaare T2-T0 und T3-T0 miteinander verschmolzen, d.h., ihre Kontrastivität geht verloren. Werden Zielwörter in Äußerungen realisiert, verschmelzen zusätzlich die Konturen von T1-T0 und T4-T0. Die Ergebnisse zeigen außerdem, dass bestimmte F0-Konturen von deutschen Lernern bei der Tonproduktion bevorzugt werden. Lerner realisieren Tonsequenzen, deren Onsets hoch sind, mit einer durchgehend fallenden Kontur über zwei Silben (in T1-T0 und T4-T0). Die Tonsequenzen mit einem tiefen Onset (in T2-T0 und T3-T0) werden mit einer steigenden Kontur in der ersten und einer fallenden in der zweiten Silbe produziert. Hinsichtlich des Dauerverhältnisses zwischen den beiden Silben weisen Lerner und Muttersprachler ähnliche Muster auf.

Des Weiteren zeigen beide Studien, dass sich mit Vergrößerung der tontragenden Domäne der Unterschied zwischen der Tonproduktion von Lernern und Muttersprachlern vergrößert.

7.1.2 Ergebnisse im Bereich tonaler Koartikulation

Die Untersuchung der vorliegenden Arbeit zur tonalen Koartikulation (Kapitel 5) basiert hauptsächlich auf der Annahme von Flemming (2011), dass diese zum einen von universellen *constraints*, wie beispielsweise der maximalen Schwingungsgeschwindigkeit der Stimmlippen, und zum anderen von sprachspezifischen Tontargets bestimmt wird. Die erste Studie der vorliegenden Arbeit zeigt, dass tonale Koartikulation auch in der L2-Produktion existiert und

somit den universellen *constraints* unterliegt. Tonale Koartikulation in der L2-Produktion zeigt allerdings ein spezifisches Muster, das sich von dem der muttersprachlichen Produktion stark unterscheidet. Zusätzlich wird herausgefunden, dass die Auswirkungen des regressiven koartikulatorischen Effektes weitreichend sind und sich nicht nur auf das Offset eines Tons beschränken. Dies entspricht dem Befund der Studie von Brunelle (2009) und deutet darauf hin, dass diese Eigenschaft keine spezielle Erscheinung innerhalb des Spracherwerbs, sondern sprachübergreifend ist. Bezüglich des Wesens der tonalen Koartikulation wird festgestellt, dass progressive tonale Koartikulation in der muttersprachlichen Produktion von assimilatorischer Natur und vorwiegend regressive tonale Koartikulation in der Lerner-Produktion dissimilatorisch ist. Die starke progressive Assimilation in der muttersprachlichen Produktion kann durch die Implementierungsregeln der Tontargets im SC plausibel erklärt werden. Jedoch bleibt bei der dominanten regressiven Dissimilation in der Lerner-Produktion unklar, ob diese durch den Unterrichtseffekt oder die Implementierung der Tontargets im Deutschen hervorgerufen wird.

In der muttersprachlichen Tonproduktion zweisilbiger Wörter mit einem Neutralton zeigt sich ebenfalls ein starker progressiver tonaler koartikulatorischer Effekt, der in der Lerner-Produktion nicht feststellbar ist. Der Tonhöhenverlauf über der Hälfte der Silbe mit einem Neutralton variiert je nach Offset-Wert des vorausgehenden vollen Tons beträchtlich. In bisherigen Studien wird kein expliziter Zusammenhang zwischen der Oberflächenkontur des Neutraltons und tonaler Koartikulation hergestellt. Verschiedene Autoren interpretieren die Konturen des Neutraltons als ein Resultat der Interpolation zwischen zwei Tontargets oder der Tonausbreitung. Die Ergebnisse der vorliegenden Arbeit unterstützen jedoch die These von Chen und Xu (2006), dass der Neutralton ein eigenes grundlegendes Tontarget besitzt, da dieses dem tonalen koartikulatorischen Effekt unterliegt.

7.2 Diskussionen der Ergebnisse

Im Folgenden werden die Ergebnisse anhand der in Kapitel 4 aufgestellten Hypothesen diskutiert. Dabei stehen insbesondere der linguistische und der didaktische Faktor im Fokus, welche die folgenden Fragen beantworten sollen:

(3) Welche Mechanismen verwenden deutsche ChaF-Lerner bei der Realisierung von Tonsequenzen?

(4) Welche Rolle spielt der bisher praktizierte ChaF-Unterricht?

7.2.1 Tonale Koartikulation in der L2-Produktion

Die Ergebnisse der ersten empirischen Studie der vorliegenden Arbeit bestätigen die Hypothesen H1 und H2, wonach deutsche Lerner bei tonaler Koartikulation in der Realisierung zweisilbiger Wörter grundlegend andere Mechanismen einsetzen als Muttersprachler.

H1: Die tonale Realisierung von Tonsequenzen durch deutsche ChaF-Lerner zeigt eine vorwiegend regressive koartikulatorische Tendenz, während chinesische Muttersprachler eine vorwiegend progressive Tendenz zeigen.

Es gibt mehrere mögliche linguistische Erklärungen für die bei der Produktion der Lerner gefundene vorwiegend regressive Koartikulation. Die tontragende Domäne spielt hierbei eine bedeutende Rolle. Wie in Kapitel 5 erläutert, werden Tontargets im SC mit dem Silbenende aligniert. Um tonale Kontraste zu gewährleisten, findet der Übergang zwischen zwei Tontargets daher zwangsläufig am Onset des nächsten Tons statt. Im Deutschen kann die tontragende Domäne hingegen vom Umfang einer Silbe bis zum Umfang einer Äußerung variieren und der Tonakzent ist in der betonten Silbe verankert. Die Alignierung des Tontargets orientiert sich am Anfang der betonten Silbe und dem der unmittelbar

nachfolgenden Silbe. Das Onset eines Tons wird von Lernern und Muttersprachlern unterschiedlich gewichtet, sodass Lerner viel weniger Variation im Onset zulassen und den Übergang entsprechend frühzeitig beenden, wie in Abbildung 5.5 (c) dargestellt. Dies könnte zu vorwiegend regressiver tonaler Koartikulation führen.

Des Weiteren spielt die Zuweisung von Wortakzenten im Deutschen eine Rolle. Während das Gewicht der beiden Silben in einem zweisilbigen Kompositum im SC nahezu gleich ist, wird der Wortakzent in einem deutschen zweisilbigen Wort einer der beiden Silben zugewiesen. Féry (1996) zeigt, dass im Deutschen die letzte Silbe eines zweisilbigen Worts den Wortakzent erhält, wenn es eine offene Silbe mit einem langen Vokal ist. Dohmas (2008) findet zudem heraus, dass deutsche Sprecher Vollvokale in einer offenen Silbe als schwere oder gespannte Vokale realisieren. In der vorliegenden Studie ist die letzte Silbe aller Zielwörter eine offene Silbe mit einem langen Vokal oder einem Diphthong. Es ist wahrscheinlich, dass Lerner der letzten langen Silbe oder dem Diphthong den Wortakzent zuweisen. Dies könnte dazu führen, dass Lerner den zweiten Ton mit mehr Aufmerksamkeit produzieren und die tonale Identität des ersten Tons zu seinen Gunsten vernachlässigen. Dieses Phänomen könnte durch die Hyperartikulations-Hypothese (Lindblom 1990; Stent et al. 2008) erklärt werden. Im Deutschen ist die trochäische Struktur zweisilbiger Wörter so stark, dass die zweite Silbe in den Zielwörtern der Studie als unerwartet und fremdartig wahrgenommen wird. Der volle Ton auf der zweiten Silbe wird von den Lernern als eine Silbe mit stärkerer Betonung fehlinterpretiert und deshalb hyperartikulierte.

Eine weitere mögliche Erklärung für die gefundene vorwiegend regressiver Koartikulation in der Lerner-Produktion ist das *tonal crowding* (Arvaniti et al. 2006). Viele Studien zu Intonationssprachen zeigen, dass der rechts stehende Nuklearakzent oder Grenzton einen rückwärts gerichteten Zeitdruck auf die Alignierung des Pränuklearakzents ausübt (Silvermann & Pierrehumbert 1990; Prieto et al. 1995). Dies ist eine Tendenz, die regressiver tonaler Koartikulation

entspricht und deutsche Lerner möglicherweise negativ bei der Realisierung unmittelbar nacheinander stehender voller Töne beeinflusst.

H2: Deutsche ChaF-Lerner versuchen, Töne in Tonsequenzen möglichst klar voneinander abzugrenzen, um die erlernte Zitierform der lexikalischen Töne beizubehalten. Demzufolge zeigen sie bei tonaler Koartikulation eine dissimilatorische Tendenz.

In Kapitel 5 wird gezeigt, dass Lerner in der Studie tatsächlich dissimilatorische Mechanismen einsetzen, wenn sie Sequenzen ausschließlich in Zitierform gelernter Töne produzieren. Wie in Kapitel 3 dargestellt, werden die lexikalischen Töne im Unterricht in der Regel nur in ihren Zitierformen vorgestellt und geübt. Einer der wichtigsten Punkte beim Tonerwerb wird den Lernern nicht bewusst gemacht: Höhe und Kontur jedes einzelnen Tons ändern sich je nach dessen tonalem Kontext. Deshalb versuchen Lerner, Töne in ihrer Zitierform wiederzugeben, wenn sie diese in Sequenzen produzieren. Um den tonalen Kontrast trotz unvermeidbarer Koartikulation zu gewährleisten, dissimilieren sie unbewusst die Töne. Dies deutet darauf hin, dass der Unterricht einen beachtlichen Einfluss auf die tonale Realisierung der Lerner hat.

Die Feststellung, dass die Tonsequenz T3-T3 von Lernern meistens mit einer zielsprachengerechten Kontur realisiert wird, stützt die Hypothese H3.

H3: Explizit gelernte Tonsequenzen, wie beispielsweise die T3-T3 Sequenz, werden von deutschen ChaF-Lernern und chinesischen Muttersprachlern ähnlich realisiert.

Die tonale Realisierung der Sequenz T3-T3 wird im Unterricht explizit als Einheit statt in isolierter Form vermittelt und geübt. Bei dieser Tonsequenz wird kein signifikanter Unterschied in den Konturen zwischen der Produktion von Lernern und Muttersprachlern gefunden (Kapitel 5, Abschnitt 5.3.5). Dieses Ergebnis

verdeutlicht weiter den Unterrichtseffekt bei der tonalen Realisierung in der L2-Produktion.

Zusammengefasst deuten die Ergebnisse darauf hin, dass eine explizite Anleitung zur Produktion von Tonsequenzen bei Lernern eine zielsprachengerechte Produktion fördert, während das Ausbleiben einer derartigen Anweisung zu den gefundenen dissimilatorischen Tendenzen führt.

7.2.2 Tonale Realisierung des Neutraltons

Die Ergebnisse der zweiten Studie zeigen, dass die produzierten Konturen bei der Tonproduktion der Lerner bei bestimmten Tonpaaren verschmelzen und bestätigen somit die Hypothesen H4:

- H4: Lerner produzieren den Neutralton im SC wie eine unbetonte Silbe im Deutschen, wodurch lexikalische Kontraste in der Tonproduktion zum Teil verloren gehen und die tonale Realisierung bestimmter Tonpaare miteinander verschmilzt.

Das bereits in Kapitel 6 (Abschnitt 6.1.2) erwähnte Phänomen, dass sowohl bei L1- als auch L2-Sprechern die tonalen Realisierungen von bestimmten Tonpaaren häufig miteinander verschmelzen, ist nicht die einzige Erklärung für das oben dargestellte Ergebnis.

Eine weitere Erklärung könnte in der Art und Weise liegen, wie die Realisierung des Neutraltons den Lernern im Unterricht vermittelt wird. Wie in Kapitel 3 beschrieben, werden Silben mit Neutralton und unbetonte Silben im Deutschen in Lehrbüchern und im Unterricht in den meisten Fällen gleichgesetzt. Eine unbetonte Silbe hat kein bestimmtes Target und wird lediglich kürzer und schwächer realisiert als eine betonte Silbe. Das Wort ist hier die Implementierungsdomäne für bestimmte tonale Spezifikationen. Das heißt, wenn T1 (H) oder T4 (HL) ein Neutralton folgt, wird er mit einer ähnlichen Kontur

produziert. Wie in Abbildung 6.2 des Kapitel 6 gezeigt, kann eine fallende Kontur (z.B. H*L- im deutschen Wort *männlich*) auf zwei unterschiedliche Arten realisiert werden, ohne ihre phonologische Spezifikation zu verändern. Der Gipfel des H* kann mit der betonten Silbe (im Beispiel „*männ-*“) früh oder spät aligniert werden. Die F0-Kontur der Sequenz T1-T0 im Wort *bēi zi* hat eine ähnliche Oberflächenkontur wie die im Wort *männlich*, in welcher der Gipfel des H* spät in der Silbe „*männ-*“ aligniert wird, während die F0-Kontur der Sequenz T4-T0 im Wort *mào zi* eine ähnliche Kontur wie die im Wort *männlich* hat, in welcher der Gipfel des H* früh in derselben Silbe „*männ-*“ aligniert wird. Der ausschlaggebende Unterschied liegt jedoch darin, dass die Konturen der Sequenzen T1-T0 und T4-T0 lexikalisch distinktiv sind. Aufgrund der fehlenden Strategien bei der Tonproduktion von Wörtern mit Neutralton im SC realisieren deutsche Lerner die von T1-T0 und T4-T0 nicht kontrastiv, sondern wählen in beiden Fällen die ihnen vertraute F0-Kontur aus ihrer L1, wie z.B. die von H*L-.

Unter allen vier Tonpaaren sind die Realisierungen von T3-T0 am weitesten von der muttersprachlichen Produktion entfernt. Die Realisierung von T3-T0 der Muttersprachler ähnelt der Kontur eines Tonakzentes L*+H im Deutschen. Diese Kontur signalisiert im Deutschen üblicherweise, dass eine Aussage als emotionale oder sogar sarkastische Behauptung gedeutet werden soll, wie in Abbildung 7.1 zu sehen ist. Der Kontext im Trägersatz dieser Studie ruft keine solche Emotion hervor. Eventuell vermeiden die deutschen Lerner bewusst die in diesem Kontext für sie unpassend erscheinende Kontur, da diese im Deutschen pragmatisch „vorbelastet“ ist.

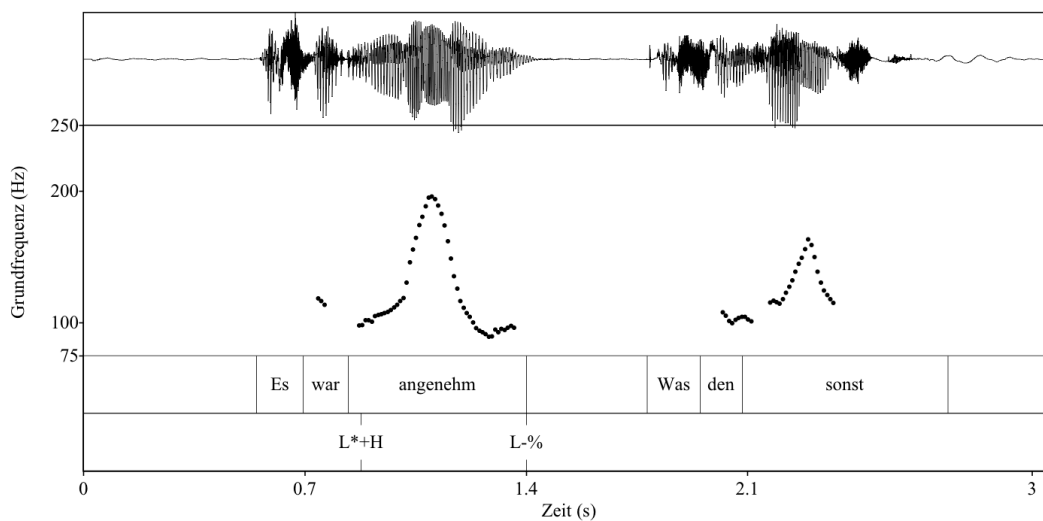


Abbildung 7.1: Beispiel für den Tonakzent L*+H (www.gtobi.uni-koeln.de)

Zum einen zeigen die betreffenden Silben in zweisilbigen Wörtern mit Neutralton im SC und in einem Trochäus im Deutschen ein ähnliches Betonungsverhältnis. Zum anderen wird eine Silbe mit Neutralton ohne jegliche explizite Übung im Unterricht mit der unbetonten Silbe im deutschen Trochäus gleichgesetzt und daher wie eine solche produziert. Aus diesen Gründen zeigen die Lerner Tendenzen, die stark auf ihre Muster in der L1 zurückzuführen sind. Dieser Einfluss der L1 wird in einer größeren Domäne besonders deutlich.

7.2.3 Vorläufiges Modell des Tonerwerbs deutscher ChaF-Lerner

Aus den Ergebnissen und Diskussionen der vorliegenden Arbeit ergibt sich ein vorläufiges Modell des Tonerwerbs deutscher ChaF-Lerner:

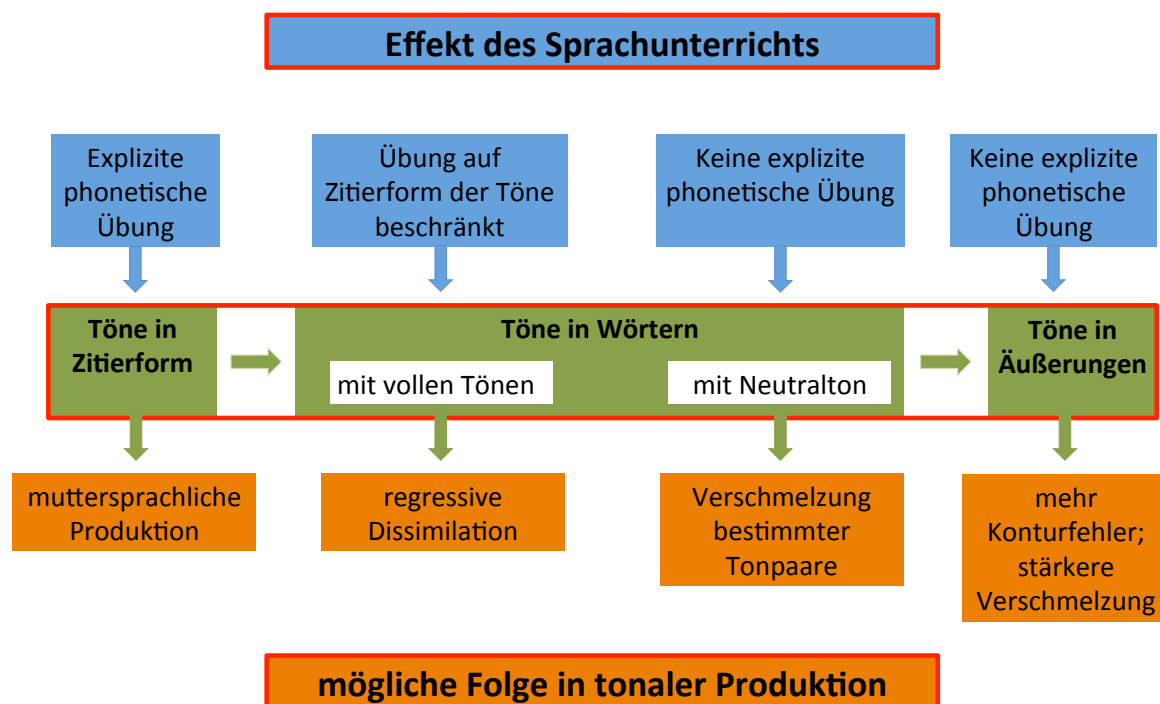


Abbildung 7.2: Vorläufiges Modell zum Tonerwerb deutscher ChaF-Lerner

In diesem Modell wird die Rolle des ChaF-Unterrichts besonders betont, da die in den Untersuchungen der vorliegenden Arbeit gefundenen Mechanismen bei der Tonproduktion der Lerner stark auf die Art und Weise zurückzuführen sind, wie Töne im ChaF-Unterricht vermittelt werden. Linguistische Faktoren inklusive des Einflusses aus der L1 werden im obigen Modell als die Folge des ChaF-Unterrichts interpretiert. Mittels dieses Modells soll klar dargestellt werden, dass eine Änderung im Vermittlungsprozess der lexikalischen Töne im ChaF-Unterricht notwendig ist.

7.3 Didaktische Konsequenzen

7.3.1 Gewichtung des Aussprachetrainings

Dass explizites Aussprachetraining einen positiven Effekt auf die Aussprachekompetenz bei Fremdsprachenlernern haben kann, wird in vielen Studien bereits bestätigt (z.B. Cenoz & Lecumberri 1996; Bongaerts et al. 1997; Elliott 1997; Derwing, Munro & Wiebe 1997; Derwing & Rossiter 2003; Dahmen 2013).

Die Arbeit von Dahmen (2013) zeigt, dass eine Gruppe italienischer Deutsch-Lerner mit einem gezielten prosodischen Training eine signifikant verbesserte Aussprache auf der suprasegmentalen sowie auf der segmentalen Ebene erreicht. Eine andere Gruppe, die im gleichen Zeitraum kein spezielles Training bekommt, zeigt keine derartige Verbesserung in der Aussprache.

Konkret für den Erwerb der Töne im SC zeigen auch Studien von Wang, Jongman & Sereno (2006), He (2010) und He & Wayland (2010), dass bereits mehr Sprachinput vom SC eine deutlich erkennbare positive Wirkung bei Lernern erzielen kann. In den genannten Studien wird kein speziell angefertigtes Trainingsmaterial eingesetzt, sondern der allgemeine Sprachinput des SC wird erhöht. Aus diesem Grund sollte das Aussprachetraining generell im Curriculum des ChaF-Unterrichts stärker gewichtet werden, damit Lerner bessere Ergebnisse beim Erwerb der Töne im SC erzielen können.

7.3.2 Expliziter Erwerb der phonetischen Realisierungen

Die Ergebnisse der beiden Studien weisen deutlich darauf hin, dass die herkömmliche Übung isolierter Töne für die Tonproduktion der Lerner nicht ausreichend ist. Im normalen Sprachgebrauch werden lexikalische Töne nur selten in ihrer Zitierform realisiert. Hintereinander auftretende Töne zeigen in der Regel eine andere Kontur und/oder Tonhöhe. Das Bewusstsein für diese Tatsache muss

von Anfang an in den Lernern geweckt werden, sodass sie in der Produktion von Tonsequenzen nicht versuchen, diese in der ihnen gewohnten Zitiertform vorzunehmen. Wie die Befunde der vorliegenden Studien zeigen, kann die Fokussierung auf Töne in Zitiertform zu einer nicht muttersprachlichen Produktion führen. Die Ausspracheübung im Unterricht sollte sich deshalb nicht ausschließlich auf die kanonischen Formen der lexikalischen Töne konzentrieren. Von der Einführung der vier tonalen Kategorien bis zur phonetischen Realisierung der Töne - vor allem in Sequenz - benötigen Lerner explizites Wissen und gezielte Übungen.

7.3.3 Kontrastive Darstellung von L1 und L2

Eine weitere Konsequenz für den zukünftigen Unterricht kann darin bestehen, dass Sprachlehrer Lernern bei bestimmten Strukturen oder Phänomenen eine kontrastive Darstellung von Mutter- und Zielsprache vorführen, um einen möglichen negativen Transfer aufzuzeigen und gezielt zu überwinden. Zum Beispiel bei der Realisierung des Neutraltons wird durch die Beschreibung in den Lehrbüchern und die Vermittlung im Unterricht suggeriert, dass Silben mit einem Neutralton im SC und unbetonte Silben im Deutschen gleichzusetzen seien. Somit ist ein negativer Transfer der beiden Phänomene vorherzusehen. In Wirklichkeit wird der Neutralton mit einem individuellen Tontarget realisiert. Eine kontrastive Darstellung der Silbe mit Neutralton und der unbetonten Silbe im Deutschen wäre hier hilfreich. Hierzu ist die Darstellung der Oberflächenkontur des Neutraltons notwendig, der nach dem jeweiligen vollen Ton produziert wird. Konkrete Beschreibungen und Anweisungen zur Realisierung des Neutraltons können wie z.B. in der in Kapitel 6 dargestellten Tabelle 6.1 oder Abbildung 6.2 eingesetzt werden.

7.3.4 Konkrete Maßnahmen im ChaF-Unterricht

Die Ergebnisse der vorliegenden Arbeit und die daraus zu ziehenden Konsequenzen unterstreichen die Notwendigkeit, bei der Vermittlung der lexikalischen Töne im ChaF-Unterricht mehr Gewicht auf die phonetische Realisierung von Tonsequenzen zu legen. Es sollten von Anfang an sämtliche 16 tonalen Kombinationen mit den vollen Tönen und vier Kombinationen der vier Töne mit dem Neutralton in zweisilbigen Wörtern trainiert werden. Man kann z.B. Minimalpaare bilden, in denen sich die Bedeutung von zwei Wörtern ausschließlich durch das Vertauschen der Töne vollständig ändert. Ziel ist dabei, dass Lerner die Produktion eines Tons in verschiedenen Positionen innerhalb tonaler Koartikulation üben, wie in Beispiel (7.1):

(7.1)	T1 + T3	vs.	T3 + T1
	qiāng shǒu		shǒu qiāng
	手 枪		枪 手
	Pistole		Gewehrschütze

Eine weitere mögliche Maßnahme betrifft die Methode des Trainings. Immer mehr Autoren machen auf die positive Auswirkung eines audio-visuellen Trainings aufmerksam (z.B. So 2003, 2006; Chun et al. 2012). Audio-visuelles Training bietet Lernern sprachlichen Input auf zwei Kanälen gleichzeitig. Insbesondere wird auf diese Weise das Bewusstsein der Lerner für die akustischen Charakteristika der Töne erhöht und dies wiederum unterstützt sie bei der Festlegung der tonalen Kategorien. In der Studie von Chun et al. (2012) wird gezeigt, dass sich die Tonproduktion der Sequenz T4-T1 eines amerikanischen Lernalters nach einer Woche audio-visuellem Training erkennbar verbessert hat (von b. zu c. im Vergleich zur muttersprachlichen Produktion in a. der Abbildung 7.3).

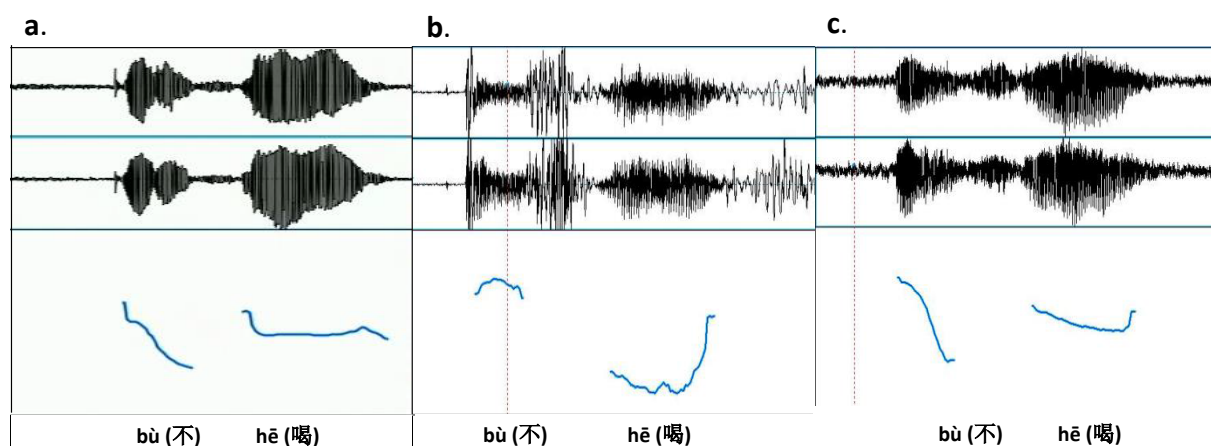


Abbildung 7.3: F0-Konturen der Wörter 不 (bù, 'nicht') und 喝 (hē, 'trinken') produziert von einem Muttersprachler (in a.) sowie einem Lerner vor (b.) und nach (c.) einem einwöchigen audio-visuellen Training.

Für die Realisierung von Wörtern mit Neutralton ist es besonders sinnvoll, den Lernern die visuellen Konturen zu zeigen, weil der Neutralton keine zugrundeliegende tonale Kategorie besitzt. Dadurch können sie das konkrete Target des Neutraltons nach den vier lexikalischen Tönen gezielter üben.

7.4 Forschungsdesiderata und Ausblick

Die vorliegende Arbeit hat erstmals tonale Koartikulation der Töne und die Realisierung zweisilbiger Wörter mit Neutralton in der L2-Produktion untersucht. Die Ergebnisse haben bestimmte grundlegende Mechanismen in der Tonproduktion offengelegt. Es hat sich gezeigt, dass die Vermittlung der lexikalischen Töne im Unterricht auch bei der tonalen Realisierung der Lerner eine Rolle spielt.

Allerdings haben sich die Untersuchungen in der vorliegenden Arbeit ausschließlich auf die Produktion der Lerner konzentriert. Wie in Kapitel 4 besprochen, hängt die Perzeption mit der Produktion der Töne eng zusammen, ohne dass eine 1:1-Relation besteht. Um ein vollständiges Bild des Tonerwerbs deutscher Lerner zu erstellen, müssten adäquate Perzeptionsexperimente durchgeführt werden. Wie in dieser Arbeit gezeigt wird, können deutsche Lerner Töne in Zitierform muttersprachähnlich produzieren. Deshalb sollte nach

Möglichkeit als erstes getestet werden, ob bei Lernern die Perzeption isolierter Töne fehlerfrei erfolgt. Die Untersuchung der Produktion von Tönen in Sequenz innerhalb der vorliegenden Arbeit hat ergeben, dass die Lerner bei tonaler Koartikulation eine den Muttersprachlern entgegengesetzte Richtung zeigen. In einer vorherigen Studie von Xu (1994) hat sich gezeigt, dass die Perzeption von Tönen stark vom tonalen Kontext abhängig ist. Durch das Wegfallen oder Manipulieren der kontextualen Information sinkt die Identifikationsquote bei den Muttersprachlern. Um zu testen, wie viel Gewicht Lerner auf die Information tonaler Koartikulation bei der Perzeption von Tönen legen, sollten Lerner anhand von Stimuli Töne mit manipulierten F0-Konturen identifizieren. Das Ergebnis könnte Aufschluss darüber geben, ob Lerner bei der tonalen Produktion und Perzeption ähnliche Mechanismen zeigen.

Die Ergebnisse der vorliegenden Arbeit zeigen die Notwendigkeit auf, den Lernern phonetische Realisierungen beim Tonerwerb im SC explizit zu vermitteln. Um zu beweisen, dass durch gezieltes Aussprachetraining die Kompetenz der Aussprache signifikant verbessert werden kann, ist es sinnvoll, eine Vergleichsstudie mit zwei Gruppen deutscher Lerner durchzuführen. Eine Gruppe wird mit der herkömmlichen Methode für Vermittlung der Töne unterrichtet. Die andere Gruppe erhält ein gezieltes Aussprachetraining. Nach einem bestimmten Zeitraum werden Produktion und Perzeption der beiden Gruppen verglichen, wobei eine klare Verbesserung der Aussprache hinsichtlich der Töne bei der speziell geschulten Gruppe zu erwarten ist.

Außerdem sollten auch drei- und viersilbige Wörter im Bezug auf tonale Koartikulation erforscht werden. Ein entscheidender Unterschied zwischen zweisilbigen und mehrsilbigen Wörtern liegt in dem Betonungsverhältnis zwischen den Silben. Während bei der ersten und zweiten Silbe in zweisilbigen Wörtern mit vollen Tönen kein deutliches Betonungsmuster zu sehen ist, wird in mehrsilbigen Wörtern die letzte Silbe am stärksten und die mittlere am schwächsten betont. Der Wortakzent als Faktor könnte also durch die

Untersuchung tonaler Koartikulation mehrsilbiger Wörter präziser bestimmt werden.

Abschließend ist anzumerken, dass noch weitere Basisforschung zum tonalen Erwerb im SC durch deutsche Lerner betrieben werden muss und Neuerungen sowie Umstrukturierungen im ChaF-Unterricht umgesetzt werden müssen, um Lernern den Erwerb des Standardchinesischen zu erleichtern.

Literaturverzeichnis

- Abramson, A. (1979). The coarticulation of tones: An acoustic study of Thai. In T. L-Thongkum, P. Kullavanijaya, V. Panupong and T. L. Tingsabadh (Hg), *Studies in Tai and Mon-Khmer Phonetics and Phonology in Honour of Eugenie J. A. Henderson*. 1-9. Bangkok, Chulalongkorn University Press.
- Alao, G. (1995). African languages abroad: problems and proposals. *Journal of the African Language Teachers Association*, 1, S.63-83.
- Anderson-Hsieh, J., Johnson, R. & Koehler, K. (1992). The relationship between native speaker judgements of non native pronunciation and deviance in segmentals, prosody and syllable structure. *Language Learning* 42, S.529-555.
- Arvaniti, A. & Ladd, D. R. (1995) Tonal alignment and the representation of accentual targets. In: K. Elenius & R. Branderad (Hg) *Proceedings of the 13th international congress of phonetic sciences*, Stockholm Vol. 4, S.220-223
- Arvaniti, A., Ladd, D. & Mennen, I. (1998). Stability of Tonal Alignment: The Case of Greek Prenuclear Accents. *Journal of Phonetics* 26, S.3-25.
- Arvaniti, A., Ladd, D. & Mennen, I. (2006). Phonetic Effects of Focus and "Tonal Crowding" in Intonation: Evidence from Greek Polar Questions. *Speech Communication* 48, S.667-696.
- Ashby, M. (1978) A study of two English nuclear tones. *Language and Speech*, 21, S. 326-336.
- Atterer, M., & Ladd, D.R. (2004). On the phonetics and phonology of "segmental anchoring" of F0: evidence from German. *Journal of Phonetics* 32, S.177-197.
- Bao, Z.M. (1990). *On the Nature of Tone*. MIT, Cambridge, Mass, Dissertation.

- Baumann, S. (2006). *The Intonation of Givenness - Evidence from German*. Saarland University, Dissertation. Linguistische Arbeiten 508. Tübingen: Niemeyer.
- Baumann, S., Grice, M. & Steindamm, S. (2006). Prosodic Marking of Focus Domains - Categorical or Gradient? *Proceedings Speech Prosody 2006*, Dresden, S.301-304
- Bausch, K. R.; Christ, H. & Krumm, H. J. (Hg) (2003). *Handbuch Fremdsprachenunterricht*. 4. Aufl. A. Francke Verlag Tübingen und Basel.
- Best, C.T. (1993). Emergence of language-specific constraints in perception of non-native speech contrasts: A window on early phonological development. In: B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage & J. Morton (Hg), *Development neurocognition: Speech and face processing in the first year of life*. S.289-304. Dordrecht: Kluwer Academic Publishers.
- Best, C.T. (1994). Learning to perceive the sound pattern of English. In: C. Rovee-Collier & L. Lipsitt (Hg). *Advance in Infancy Research* 9. S.217-304, Hillsdale, NJ: Ablex Publishers.
- Best, C.T. & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In: O-S. Bohn, M. Munro (Hg), *Language Experience in Second Language Speech Learning. In honor of James Emil Flege*. S.13-34, Amsterdam: John Benjamins.
- Bent, T. (2005). *Perception and Production of Non-native Prosodic Categories*. Northwestern University, Dissertation.
- Bladon, A. & Al-Bamerni, A. (1976). Coarticulation resistance in English /l/. *Journal of Phonetics* 4. S.137-150.

- Bluhme, H & Burr, R.(1971). An audio-visual display of pitch for teaching Chinese tone. *Studies in Linguistics* 22, S.51-57.
- Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer (Version 5.3.48) [Computer software]. Retrieved from <http://www.praat.org>.
- Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition* 14, S.131-158.
- Bongaerts, T., van Summeren, C., Planken, B. & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition* 19, S.447-465.
- Bulletin of HSK Test (2011). <http://www.chinesetest.cn>
- Brunelle, M. (2003). Tone Coarticulation in Northern Vietnamese. *Proceedings of the 15th International Congress of Phonetic Sciences*. S.2673-2676.
- Brunelle, M. (2009) Northern and Southern Vietnamese Tone Coarticulation: A Comparative Case Study. *Journal of Southeast Asian Linguistics* 1, S.49-62.
- Cenoz, J., & Lecumberri, L. G. (1999). The effect of training on the discrimination of English vowels. *International Review of Applied Linguistics* 37, S.261-275.
- Chang, Y.C., Hsieh, F.F. (2012). Tonal coarticulation in Malaysian Hokkien: a typological anomaly? *The Linguistic Review* 29, S.37-73.
- Chao, Y.R. (1930). A system of tone letters. *La Maitre Phonétique*, 30, S.24-27.
(Reprinted in English orthography in *Fangyan*, 1980, 2, S.81-83.)

- Chao, Y.R. (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chen, M. (2000). *Tone Sandhi: Patterns across Chinese Dialects*. Cambridge Studies in Linguistics 92. Cambridge: Cambridge University Press.
- Chen, Q. (1997). Toward a sequential approach for tonal error analysis. *Journal of Chinese Language Teachers Association* 32, 1, S.21-39.
- Chen, Q. (2000). *An Analysis of Mandarin Tonal Errors in Connected Speech by English-speaking American Adult Learners: A Study at and above the Word Level*. Brigham Young University, Dissertation.
- Chen, Y. (2003). *The Phonetics and Phonology of Contrastive Focus in Standard Chinese*. State University of New York at Stony Brook. Dissertation.
- Chen, Y., & Braun, B. (2006). Prosodic realization in information structure categories in Standard Chinese. In: R. Hoffmann, & H. Mixdorff (Hg) *Speech Prosody 2006*. Dresden: TUD Press.
- Chen, Y., & Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics* 36, S.724-746.
- Chen, Y. & Xu, Y. (2006). Production of weak elements in speech - Evidence from f0 patterns of neutral tone in standard Chinese. *Phonetica* 63, S.47-75.
- Chen, Y. (2010). Post-focus F0 compression - Now you see it, now you don't. *Journal of Phonetics* 38, S.517-525
- Cheng, C. (1968). English stresses and Chinese tones in Chinese sentences. *Phonetic* 18, S.77- 88.

- Cheng, C. (1973). *A Synchronic Phonology of Mandarin Chinese*. Monographs on Linguistic Analysis 4. The Hague: Mouton.
- Cho, T. (2002). *The Effects of Prosody on Articulation in English*. Routledge. New York, NY.
- Cho, T.(2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics* 32 , S.141-176
- Chun, D. M., Jiang, Y., & Ávila, N. (2013). Visualization of tone for learning Mandarin Chinese. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*. Aug. 2012. S. 77-89. Ames, IA: Iowa State University.
- Cremerius, R. (2004). *Chinesisch für Deutsche 1*. Hamburg: Helmut Buske Verlag GmbH.
- Dahmen, S. (2013). *Prosodie oder Segmente? Phonetische Untersuchungen zu Trainingseffekten bei italienischen Deutschlernenden*. Universität zu Köln, Dissertation. Online-Publikation, KUPS.
- Derwing, T. M., Munro, M. J., & Wiebe, G. E. (1997). Pronunciation instruction for fossilized learners: Can it help? *Applied Language Learning* 8, S.217-235.
- Derwing, T.M. & Rossiter, M.J. (2003). The effects of pronunciation instruction on the accuracy, fluency and complexity of L2 accented speech. *Applied Language Learning*, 13, S.1-18.
- DiCanio, C. (2014) Triqui tonal coarticulation and contrast preservation in tonal Phonology. *Proceedings from Sound Systems of Mexico and Central America*, Yale University, New Haven, CT.

- Ding, H (2012). Perception and Production of Mandarin Disyllabic Tones by German Learners. *Speech Prosody*. Shanghai, China, Mai 22.-25. S.378-381
- Ding, H., Hoffmann, R., Hussein, H., Jokisch, O. (2012). An Acoustic Investigation of Chinese Tone 4 Produced by German Learners. *2012 International Conference on Asian Language Processing*. Hanoi, Vietnam, November 13.-15. S.5-8
- Domahs, U., Plag, I. & Carroll, R. (2014). Word stress assignment in German, English and Dutch: Quantity-sensitivity and extrametricality revisited. *Journal of Comparative Germanic Linguistics* 17, S.59–96.
- Duanmu, S. (1990). *A Formal Study of Syllable, Tone, Stress and Domain in Chinese Languages*. MIT, Cambridge, Mass, Dissertation.
- Duanmu, S. (2000). *The Phonology of Standard Chinese*. Oxford: Oxford University Press.
- Duanmu, S. (2004). Tone and Non-tone Languages: An Alternative to Language Typology and Parameters. *Language and Linguistics* 5.4, S.891-923.
- Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning* 27, S.315–330.
- Edmondson, W. & House, J. (2006, 3. Auflage). *Einführung in die Sprachlehrforschung*. Tübingen, Basel: Francke (UTB).
- Elliot, C. E. (1991). *The relationship between the perception and production of Mandarin tones: An exploratory study*. University of Hawai'i Working Papers in ESL. 10(2), S.177-204.

- Elliott, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania* 80, S.95–108.
- Farnetani, E. (1999). Coarticulation and Connected Speech Processes. In: W. Hardcastle and J. Laver (eds.). *The Handbook of Phonetic Sciences*. S.371-404. Malden, Blackwell.
- Farnetani, E. and D. Recasens (1999). Coarticulation models in recent speech Production theories. In: W. Hardcastle and N. Hewlett (eds.), *Coarticulation: Theory, Data and Techniques*. S.31-65. New York, Cambridge University Press.
- Féry, C. (1998). German word stress in Optimality Theory. *Journal of Comparative Germanic Linguistics* 2, S.101-142.
- Flege, J. E. (1988) Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America* 84, S.70-79.
- Flege, J. E. & Fletcher, K. L. (1992) Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America* 91, S.370-389.
- Flege, J. E., Munro, M. J. & MacKay, I. R. A. (1995) Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America* 97, S.3125-3134.
- Flege, J.E. (1995). Second language speech learning. Theory, findings and problems. In: Strange, Winifred (Hg). *Speech perception and linguistic experience. Issues in cross-language research*. Timonium, MD: York Press, S. 233-77.

- Flege, J. E. (1999). Age of learning and second language speech' in Birdsong D. (Hg.): *Second Language Acquisition and the Critical Period Hypothesis*. Lawrence Erlbaum, S.101-32
- Flemming, E. (2011). The grammar of coarticulation (English draft). In: M. Embarki and C. Dodane (Hg.) *La Coarticulation: Indices, Direction et Representation*.
- Gandour, J., S. Potisuk and S. Dechongkit (1992a). Anticipatory tonal coarticulation in Thai noun compounds. *Linguistics of the Tibeto-Burman Area* 15, S.111-124.
- Gandour, J., S. Potisuk and S. Dechongkit (1992b). Tonal coarticulation in Thai disyllabic utterances: A preliminary study. *Linguistics of the Tibeto-Burman Area* 15, S.93-110.
- Gandour, J., S. Potisuk and S. Dechongkit (1994). Tonal Coarticulation in Thai. *Journal of Phonetics* 22(4), S.477-492.
- Gårding, E. (1987). Speech act and tona pattern in standard Chinese – constancy and variation. *Phonetica* 44, S.13-29.
- Giegerich, H. (1985). *Metrical phonology and phonological structure: German and English*. Cambridge: Cambridge University Press.
- Gottfried, T. L., & Suiter, T. L. (1997). Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones. *Journal of Phonetics*, 25, S.207-231.
- Grice, M., R. Benz Müller, M. Savino & B. Andreeva (1995). The Intonation of queries and checks across languages: Data from Map Task Dialogues.

Proceedings of the XIII International Congress of Phonetic Sciences S.648-651.
Stockholm, Sweden.

Grice, M., Reyelt, M., Benzmüller, R., Mayer, J. & Batliner, A. (1996). Consistency in Transcription and Labelling of German Intonation with GToBI. *Proceedings of the Fourth International Conference on Spoken Language Processing*, Philadelphia, S.1716-1719.

Grice, M. & S. Baumann (2002). Deutsche Intonation und GToBI. *Linguistische Berichte* 191, S.267-298.

Grice, M., Baumann, S. & Benzmüller, R. (2005). German Intonation in Autosegmental- Metrical Phonology. In S.A. Jun (Hg.). *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

Grice, M. (2006). Intonation. In K. Brown (Hg.). *Encyclopedia of Language and Linguistics*, 2. Auflage, vol 5. S.778-788, Oxford: Elsevier.

Grice, M. & S. Baumann (2007). An Introduction to Intonation – Functions and Models. In: Trouvain, Jürgen & Ulrike Gut (Hg.). *Non-Native Prosody. Phonetic Description and Teaching Practice*. Berlin, New York: De Gruyter, S. 25-51

Guder, A.(2005). Chinesisch und der Europäische Referenzrahmen. Einige Beobachtungen zur Erreichbarkeit fremdsprachlicher Kompetenz(en) im Chinesischen. In: *CHUN - Chinesischunterricht* 20, S.83-98.

Halle, M. & Idsardi, W. (1995). General properties of stress and metrical structure. In: John Goldsmith (Hg.). *The Handbook of Phonological Theory*. S.403-443. Cambridge, Mass.: Blackwell.

- Halle, M.(1998). The stress of English Words: 1968–98. *Linguistic Inquiry* 29.4, S. 539 – 568.
- Han, M. and K.-O. Kim (1974). Phonetic variation of Vietnamese tones in disyllabic utterances. *Journal of Phonetics* 2, S.223-232.
- Hao, Y-C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and nontonal language speakers. *Journal of Phonetics* 40, S.269–279.
- Hayse, B.(1995). *Metrical Stress Theory: Principles and case studies*. Chicago: University of Chicago Press.
- He, Y. J., Wang, Q., & Wilshire (2008). Production of English lexical stress by inexperienced and experienced learners of English. *The annual conference of Canadian Acoustical Association*. Vancouver, Canada.
- He, Y. (2010). *Perception and Production of Isolated and Coarticulated Mandarin Tones by American Learners*. University of Florida, Dissertation.
- He, Y. and Wayland, R. (2010). Production of Mandarin coarticulated tones by inexperienced and experienced English learners of Mandarin. *Proceedings of the 5th International Conference on Speech Prosody*, 100123. S.1-4.
- Henrici, G., Vollmer, H.J. et al. (2001). Lernen und Lehren von Fremdsprachen: Kognition, Affektion, Interaktion. Ein Forschungsüberblick. *ZFF* 12 (2), S.1-145.
- Hsiao, Y. (1991). *Syntax, rhythm and tone: a triangular relationship*. University of California, San Diego, Dissertation.
- Ho, A.T. (1976). The acousutic variation of Mandarin tones. *Phonetica* 33, S.353-367

- Howie, J.M. (1976). *Acoustical studies of Mandarin vowels and tones*. New York: Cambridge University Press.
- Hung, T. (1987). *Syntactic and semantic aspects of Chinese tone sandhi*. University of California, San Diego, Dissertation.
- Janssen [Domahs], U. (2003). *Untersuchungen zum Wortakzent im Deutschen und Niederländischen*. Düsseldorf: University of Düsseldorf, Dissertation.
- Johansson, S. (1978). *Studies in error gravity*. Gothenburg: Gothenburg University.
- Jun, S.-A. (ed.) (2004). *Prosodic Typology. The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.
- Kaan, E., Barkley, C.M., Bao, M., Wayland, R. (2008). Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: an event-related potentials training study, *BMC Neuroscience* 9, 53.
- Keating, P. (1990). The window model of coarticulation: articulatory evidence. In: M. Beckman and J. Kingston (Hg.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press, S.451-470.
- Kelz, H. & Ye-Gerke, J. (2009). *Lóng – Aussprachtraining. Chinesisch A1/A2*. Stuttgart: Ernst Klett Sprachen GmbH.
- Kirilloff, C. (1969) On the auditory discrimination of tones in Mandarin. *Phonetica* 20, S.63 - 67.
- Kleber, Felicitas (2011): *Incomplete neutralization and maintenance of phonological contrasts in varieties of Standard German*. LMU München: Fakultät für Sprach- und Literaturwissenschaften, Dissertation.

- Kniffka, G. & Siebert-Ott, G. (2007). *Deutsch als Zweitsprache: Lehren und Lernen*. (Reihe Standard Wissen Lehramt). Paderborn: Schöningh
- Kozhevnikov, A. & Chistovich, L. (1965). Speech, articulation, and perception. *Washington: National Technical Research Service* 30 (US. Department of Commerce).
- Kratochvil, P. (1968). *The Chinese Language Today: Features of An Emerging Standard*. London: Hutchinson and Company.
- Kratochvil, P. (1987). The case of the third tone. In: *Wang Li Memorial Volumes*. S. 253-276.
- Kühnert, B. & Nolan, F. (1999). The origin of coarticulation. In: W. J. Hardcastle & N. Hewlett, (Hg.) *Coarticulation: Theory, Data and Techniques in Speech Production*. Cambridge University Press, Cambridge, S.7-30.
- Kuhl, P.K. (1992). Infants' perception and representation of speech: Development of a new theory. In: J. Ohala, T. Neary, B. Derwing, M. Hodge & G. Wiebe (Hg.), *Percedings of the Second Langauge International Conference on Spoken Language Processing*. S.449-456, Edmonton, Canada: University of Alberta Press.
- Kuhl, P.K. & Iverson, P. (1995). Linguistic experience and the „peceptuel magnet effect“. In: W. Strange (Hg), *Speech perception and linguistic experirence: Issues in cross-language research*. S.121-154, Timonium, MD: York Press.
- Ladd, D.R. (2008). *Intonational Phonology* (2. Auflage). Cambridge: Cambridge University Press.
- Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather & A. James (Hg.), *New Sounds*

90: *Proceedings of the 1990 Amsterdam symposium on the acquisition of second language speech*. University of Amsterdam. S.72-97.

Lenneberg, E. H. (1967) *Biological foundations of language*. New York: Wiley.

Li, C.N. & Thompson, S.A. (1977). The acquisition of tone in Mandarin speaking-children. *Journal of Child Language* 4, S.185-99.

Li, W.M. (1981). 试论轻声和重音 [A preliminary discussion on Stressless and stressed syllables]. In: *中国语文 [Chinese Phonology]*, 1981.1, S.35-40.

Li, Z. (2003). *The phonetics and phonology of tone mapping in a constraint-based approach*. MIT, Cambridge Mass., Dissertation.

Lin, H. (2006). Mandarin Neutral Tone as a Phonologically Low Tone. *Journal of Chinese Language and Computing* 16 (2), S.121-134.

Lin, T. (2001). 普通话的语音和语音教学—普通话语音标准与北京话. In: 林焘语言 学论文集》. Beijing: Commercial Press.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H & H theory. *Speech Production and Speech Modeling* 55, S.403-439.

Liu, F. and Y. Xu (2005). "Parallel encoding of focus and interrogative meaning in Mandarin intonation." *Phonetica* 62, S.70-87.

Liu, X. (Hg)(2007). *Das Neue Praktische Chinesisch I: Lehrbuch*. [新实用汉语课本]. Beijing: Beijing Language and Culture University Press.

Liu, Y. (Hg)(2005). *Chinesisch Erleben – Leben in China*. [体验汉语—生活篇]. Beijing: Higher Education Press.

- Long, M. H. (1990) Maturational constraints on language development. *Studies in Second Language Acquisition* 12, S.251-285.
- Ma, X.B & Y. Zhao (Hg) (2002). 普通话教程. Guangzhou: Jinan University Press.
- Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Manuel, S.Y. 1990. The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America* 88, S.1286-1298.
- Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics* 32, S.543-563
- Mennen, I. (2007). Phonological and phonetic influences in non-native intonation. In J. Trouvain & U. Gut (Hg.), *Non-native prosody* (pp. 53-76). Berlin-New York: Mouton de Gruyter.
- Miracle, W. C. (1989). Tone production of American students of Chinese: A preliminary acoustic study. *Journal of Chinese Language Teachers Association* 24, 3, S.163-176.
- Missaglia, F. (2001). Neue Kontrastivität: Die prosodische Wende. In: Kuri, S. & Saxer, R. (Hg.), *Deutsch als Fremdsprache an der Schwelle zum 21. Jahrhundert. Zukunftsorientierte Konzepte und Projekte*. Innsbruck: Studienverlag, S.75-95.
- Munro, M. J. & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech* 38, S.289-306.

- Nash, R. (1972). Phonemic and prosodic interference and their effects on intelligibility. *Proceedings of the Seventh International Congress of Phonetic Sciences*, S.570-573.
- Orie, O. O. (2006). L2 Acquisition and Yoruba Tones: Issues and Challenges. In F. Olaoba, M. Arasanyin & M. A. Pemberton (Hg.). *Selected Proceedings of the 36th Annual Conference on African Linguistics*, S.121-128. Somerville, MA: Cascadilla Proceedings Project.
- Orten, J. (2013). Developing Chinese oral skills – A research base for practice. In: I. Kecskes (Hg.), *Research in Chinese as a Second Language*, S.9-31, Boston-Berlin: Walter de Gruyter.
- Palmer, A. (1969). Thai tone variants and language teacher. *Language Learning* 19, S.287-299
- Peng, S.H. (1996). *Phonetic Implementation and Perception of Place Coarticulation and Tone Sandhi*. Ohio State University, Dissertation.
- Peng, S.H. (1997). Production and perception in Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics* 25, S.371-400.
- Peng, S. H., Chan, M. K. M. Tseng, C. Y., Huang, T., Lee, O. J. & Beckman, M. E. (2005). Toward a Pan-Mandarin System for Prosodic Transcription. In Sun, S.,(Hg.) *Prosodic Typology: The Phonology of Intonation and Phrasing*, S.230-270. Oxford: Oxford University Press.
- Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*. MA: MIT Press, Dissertation.
- Pierrehumbert, J. & Beckman, M. (1988). *Japanese Tone Structure*. Cambridge, MA: MIT Press.

- Piske, T., MacKay, I. & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29, S.191-215.
- Port, R., Mitleb, F., & O'Dell, M. (1981). Neutralization of obstruent voicing in German is incomplete. *Journal of the Acoustical Society of America* 70, S.S13, F10.
- Potisuk, S., Gandour, J. & Harper, M. P. (1996). Contextual variations in trisyllabic sequences of Thai tones. *Phonetica* 53. S.200 – 220
- Prieto, P., van Santen, J. & Hirschberg, J. (1995) Tonal alignment patterns in Spanish. *Journal of Phonetics* 23, S.429-451
- Reyelt, M., Grice, M., Benzmuüller, R., Mayer, J. and Batliner, A. (1996). Prosodische Etikettierung des Deutschen mit ToBI. In: D. Gibbon (Hg.), *Natural Language and Speech Technology, Results of the third KONVENS conference* S.144-155. Berlin, New York: Mouton de Gruyter.
- Rösler, D. (1994). *Deutsch als Fremdsprache*. Stuttgart, Weimar: Metzler.
- Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning* 19, S.245-253.
- Scovel, T. (1988). *A time to speak: a psycholinguistic inquiry into the critical period for human speech*. New York: Newbury House/Harper & Row.
- Selinker, L. (1972). Interlanguage. In: *General Linguistics* 9/2, S.671 – 92.
- Shen, X. S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of Chinese Language Teachers Association* 24, 3, S.27-47.

- Shen, X.S. (1990a). *The Prosody of Mandarin Chinese*. Berkeley: University of California Press.
- Shen, X.S. (1990b). Tonal Coarticulation in Mandarin. *Journal of Phonetics* 18, S.281-295
- Shen, X.S. (1992). Mandarin neutral tone revisited. *Acta Linguist Hafniensia* 24, S.131-151.
- Shi, F. (1986). The bi-syllabic tone in Tianjin dialect. *Language Study* 1, S.71-90.
- Shih, C. (1986). *The prosodic domain of tone sandhi in Chinese*. University of California, San Diego, Dissertation.
- Shih, C. (1987). *The phonetics of the Chinese tonal system*. AT&T Bell Labs, Murray Hill.
- Shih, C. (1988) Tone and intonation in Mandarin. *Working Papers. Cornell Phonetics Laboratory* 3, S.83-109.
- Shih, C. & Lu, H. D. (2010). Prosody Transfer and Suppression: Stages of Tone Acquisition. *Proceedings of Speech Prosody*, Chicago, May 11-14.
- Silverman, K. & Pierrehumbert, J. (1990). The timing of prenuclear high accents in English. In: J. Kingston & M. Beckman (Hg.) *Papers in laboratory phonology I*. S.72-106. Cambridge: Cambridge University Press
- So, C. K.L. (2003). Training non-native listeners to acquire Mandarin tones with visual and auditory feedback. *Paper presented at the World CALL conference*, Banff, Alberta, Canada.

- So, C. K. L. (2006). *Effects of L1 Prosodic Background and AV training on Learning Mandarin Tones by Speakers of Cantonese, Japanese, and English*. Simon Fraser University, Dissertation.
- Stent, A. J., Huffman, M. K., Brennan, S. E. (2008). Adapting speaking after evidence of misrecognition: Local and global hyperarticulation. *Speech Communication* 50(3), S.163–178
- Sun, S. H. (1998). *The Development of a Lexical Tone Phonology in American Adult Learners of Standard Mandarin Chinese*. University of Hawaii Press, Dissertation.
- Terken J. & Hermes. D.J.(2000). The perception of prosodic prominence. In: M. Horne (Hg.) *Prosody: Theory and Experiment, studies presented to Gösta Bruce*, S.89-127, Kluwer Academic Publishers, Dordrecht.
- Trofimovitch, P. & Baker, W. (2006). Learning second language suprasegmentals. Effects of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in second language acquisition* 28-1, S.1-30.
- Trubetzkoy, N.S. (1939). Grundzüge der Phonologie. In: *Travaux du cercle linguistique de Prague* 7, S.41-50.
- Tseng, C. Y. (1981). *An acoustic phonetic study on tones in Mandarin Chinese*. Brown University, Dissertation.
- Uhmann, S. (1991). *Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie*. Tübingen: Niemeyer.
- van Santen, J.; Shih, C.; Möbius, B. (1998). Intonation. In: *Sproat, Multilingual text-to-speech synthesis: the Bell Labs approach*, S.141–190. Dordrecht: Kluwer Academic Publishers.

- Vennemann, T. (1990). Syllable structure and simplex accent in Modern Standard German. *Chicago Linguistic Society* 26(2), S.399-412.
- Wang, B. and Y. Xu (2006). Prosodic encoding of topic and focus in Mandarin. *Speech Prosody*. Dresden.
- Wang, Y., A. Jongman, & J. Sereno. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America*. 113 (2), S.1033-1043.
- Whalen, D. H. (1990). Coarticulation is largely planned. *Journal of Phonetics* 18, S.3-35.
- White, C. (1981). Tonal pronunciation errors and interference from English intonation. *Journal of Chinese Language Teachers Association* 16, 2, S.27-56.
- Wiese, R (1988). *Silbische und lexikalische Phonologie – Studien zum Chinesischen und Deutschen*. Tübingen: Max Niemeyer Verlag
- Wiese, R. (2000). *The phonology of German*. Oxford: Oxford University Press.
- Winter, B & Röttger, T. (2011). The Nature of Incomplete Neutralization in German: Implications for Laboratory Phonology. *Grazer Linguistische Studien* 76, S.55-74.
- Wu, Z. (1982). 普通话语句中的声调变化 [Pitch changes in the connected speech of Standard Chinese]. 中国语文 [Chinese Phonology], S.439-450
- Wu, Z. (1984). 普通话三字组变调规律 [Rules of tone sandhi in trisyllabic words in Standard Chinese], 中国语言学报 [Bulletin of Chinese Linguistics] 2, S.70-92.

- Wu , Z . (1988). Tone-sandhi patterns of quadra-syllabic combinations in Standard Chinese. *Report of Phonetic Research*. Institute of Linguistics (CASS), Beijing, China, PL-ARPR / 1988, S.1 – 13
- Xu, L. (2004). Manifestation of informational focus. *Lingua* 114(3), S.277-299.
- Xu, Y. and Sun X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America* 111, S.1399-1413
- Xu, Y. (1994). Production and Perception of Coarticulated Tones. *Journal of the Acoustical Society of America* 95(4), S.2240-2253.
- Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics* 25, S.61-83.
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics* 27, S.55-105
- Xu, Y. and Wang, Q. E. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication* 33, S.319-337
- Xu, Y. (2005). Speech melody as articulatorily implemented communicative functions. *Speech Communication* 46, S.220-251
- Xu, Y. (2013). ProsodyPro – A Tool for Large-scale Systematic Prosody Analysis. In: *Proceedings of Tools and Resources for the Analysis of Speech Prosody* (TRASP 2013), Aix-en-Provence, France. 7-10.
- Xu, Y. (2011). Post-focus compression: Cross-linguistic distribution and historical origin. *Proceedings of 17th International Congress of Phonetic Sciences*, Hong Kong, S.152-155.

- Yang, B. (2010). *A model of Mandarin tone categories - a study of perception and Production*. University of Iowa, Dissertation.
- Yang, C. (2011). *The Acquisition of Mandarin Prosody by American Learners of Chinese as a Foreign Language (CFL)*. Ohio University, Dissertation.
- Yang, J. (Hg)(2006). *Hanyu Jiaocheng I*. [汉语教程 第一册]. Beijing: Beijing Language and Culture University Press.
- Yip, M. (1980). *The tonal phonology of Chinese*. MIT, Cambridge, Dissertation.
- Yip, M. (1989). Contour Tones. *Phonology* 6, S.149-74.
- Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.
- Zhang, H. (2007). *A Phonological Study of Second Language Acquisition of Mandarin Chinese Tones*. The University of North Carolina at Chapel Hill, Magisterarbeit.
- Zhang, H. (2010). Phonological universals and tone acquisition. *Journal of the Chinese Language Teachers Association* 45, 1, S.39-65.
- Zhang, J. (2004). The role of contrast-specific and language-specific phonetics in contour tone distribution, B. Hayes, R. Kirchner & D. Steriade (Hg.) *Phonetically-Based Phonology*, S.157-190, Cambridge: Cambridge University Press.
- Zhang, Z. (1988). *Tone and tone sandhi in Chinese*. Ohio State University, Columbus, Dissertation.

Appendix A: Fragebogen für deutsche Probanden

Persönliche Informationen:

Name: _____ Alter: _____

Geburtsort: _____

1. Ist Ihre Muttersprache Deutsch (Muttersprache = erste erworbene Sprache im Leben)?

Ja ☐ Nein ☐

2. Sprechen Sie einen deutschen Dialekt? Wenn ja, welchen: _____

3. Wie lange studieren Sie schon Chinesisch? _____ Semester

4. Wie alt waren Sie, als Sie den ersten Chinesisch-Kurs hatten? _____

5. Bitte sortieren Sie die vier lexikalischen Töne nach der Schwierigkeit **FÜR SIE**

PERSÖNLICH:

_____ _____ _____ _____
(der schwierigste) (der einfachste)

6. Wie viel Zeit sprechen Sie außerhalb der Uni durchschnittlich in der Woche

Chinesisch? _____ Min.

7. Wie oft waren Sie in China? _____

8. Was haben Sie in China gemacht? Geben Sie bitte die Dauer Ihres Aufenthaltes an:

Reisen _____ Freunde besuchen _____

Studieren _____ Anderes: _____ Dauer: _____

9. Wo waren Sie hauptsächlich während des Aufenthaltes in China?

10. Welche Fremdsprachen haben Sie außer Chinesisch noch gelernt:

Appendix B: Vollständige Wortliste der ersten Studie

wēi 微	wéi 围	wū 乌	mī 眯	wā 蛙	mō 摸
wěi 伟	wèi 位	wú 无	mí 谜	wá 娃	mó 膜
māo 猫	máo 毛	wǔ 五	mǐ 米	wǎ 瓦	mǒ 抹
mǎo 卯	mào 帽	wù 物	mì 密	wà 袜	mò 墨

wēi māo 微猫	wēi máo 微毛	wěi māo 伟猫	wěi máo 伟毛	wū mō 乌摸	wū mó 乌膜
wēi mǎo 微卯	wēi mào 微帽	wěi mǎo 伟卯	wěi mào 伟帽	wū mǒ 乌抹	wū mò 乌墨
wéi māo 围猫	wéi máo 围毛	wèi māo 位猫	wèi máo 位毛	wú mō 无摸	wú mó 无膜
wéi mǎo 围卯	wéi mào 围帽	wèi mǎo 位卯	wèi mào 位帽	wú mǒ 无抹	wú mò 无墨

wǔ mō 五摸	wǔ mó 五膜	wā mī 蛙眯	wā mí 娃谜	wǎ mī 瓦眯	wǎ mí 瓦谜
wǔ mǒ 五抹	wǔ mò 五墨	wā mǐ 蛙米	wā mì 娃蜜	wǎ mǐ 瓦米	wǎ mì 瓦蜜
wù mō 务摸	wù mó 务膜	wá mī 娃眯	wá mí 娃谜	wà mī 袜眯	wà mí 袜谜
wù mǒ 务抹	wù mò 务墨	wá mǐ 娃米	wá mì 娃蜜	wà mǐ 袜米	wà mì 袜蜜

**Appendix C: Dialog mit eingebauten Trägersätzen zur Aufnahme der
zweisilbigen Zielwörter mit vollen Tönen**

Nǐ hǎo

A: 你 好!

Nǐ hǎo

B: 你 好!

Nǐ zuó tiān qù nǎr le

A: 你 昨 天 去 哪 儿 了?

Wǒ zuó tiān qù le _____ de miào huì

B: 我 昨 天 去 了 _____ 的 庙 会。

nǎr? _____ de miào huì?

A: 哪 儿? _____ 的 庙 会?

Bú shì, shì qù le _____ de miào huì.

B: 不 是, 是 去 了 _____ 的 庙 会。

Zhēn de ma? Hǎo wán ma?

A: 真 的 吗? 好 玩 吗?

Appendix D: Konturen sämtlicher Realisierungen der DL und CM

